

# Wavelet Characterizations for Anisotropic Besov Spaces

Reinhard Hochmuth

*FB Mathematik/Informatik, Universität Gesamthochschule Kassel, Heinrich-Plett-Strasse 40,  
34132 Kassel, Germany*

E-mail: hochmuth@mathematik.uni-kassel.de

*Communicated by Peter Oswald*

Received May 22, 2000; revised July 10, 2001

---

The goal of this paper is to provide wavelet characterizations for anisotropic Besov spaces. Depending on the anisotropy, appropriate biorthogonal tensor product bases are introduced and Jackson and Bernstein estimates are proved for two-parameter families of finite-dimensional spaces. These estimates lead to characterizations for anisotropic Besov spaces by anisotropy-dependent linear approximation spaces and lead further on to interpolation and embedding results. Finally, wavelet characterizations for anisotropic Besov spaces with respect to  $L_p$ -spaces with  $0 < p < \infty$  are derived. © 2002 Elsevier Science (USA)

*Key Words:* wavelets; anisotropic function spaces; Besov spaces; approximation spaces; Jackson estimates; interpolation; embedding.

---

## 1. INTRODUCTION

Wavelet characterizations for *isotropic* Besov spaces are part of the celebrated results in wavelet theory; see, e.g., [4–7, 10, 24, 25, 27]. In [1] wavelet characterizations for functions in *anisotropic* Besov spaces, defined via the Fourier transform, are considered by Berkolaiko and Novikov. Their approach relies basically on Fourier-analytic techniques, and as a consequence they obtain results for anisotropic spaces on  $\mathbb{R}^n$  and for smooth first-generation wavelets, in particular, for so-called David–Meyer wavelets. In [17] Garrigós and Tabacco gave characterizations for a certain subclass of anisotropic Besov spaces with respect to  $L_p$ -spaces with  $p \geq 1$ . Here we overcome the limitations of both papers. By applying systematically an approximation theoretical approach via Jackson and Bernstein estimates to two-parameter sequences of subspaces our approach covers more general wavelet bases than does [1]; e.g., biorthogonal B-spline wavelets with compact support are covered. Furthermore, our approach covers the complete possible class of anisotropic Besov spaces with respect to  $L_p$ -spaces for  $0 < p < \infty$ .

Our presentation is essentially based on ideas and methods developed for isotropic Besov spaces and given in [12, 13]. However, because of the anisotropy, several basic

properties applied in the isotropic setting do not hold or are different. Consequently, a series of arguments in the subsequent analysis have to be adapted, replaced, and generalized. Moreover, our approach includes interpolation and embedding results for anisotropic Besov spaces which seem to be new.

Anisotropic Besov spaces are a generalization of anisotropic Sobolev spaces, which contain  $L_2$ -functions possessing directional  $L_2$ -derivatives of certain orders. Thus, anisotropic function spaces are the appropriate setting for describing and analyzing functions with different smoothness properties with respect to different directions. They allow one to study systematically linear anisotropic but coercive partial differential problems and also to investigate nonlinear anisotropic boundary value problems; see [20].

Another application is presented in [21]: Studying nonlinear hyperbolic wavelet approximations shows that hyperbolic wavelets are in particular appropriate for approximating functions with anisotropic smoothness properties. That is, for certain ranges of anisotropy parameters hyperbolic wavelets approximate nonlinear functions from anisotropic Besov spaces as well as specific anisotropy adapted bases.

For more general information about anisotropic function spaces we refer the reader to [26, 29, 32, 33]. Embeddings and interpolation for anisotropic spaces can be found in [16, 30]. Wavelet-type characterizations are studied in [15, 22, 30]. We note that further generalizations with respect to moduli of smoothness for coordinatewise different  $L_p$  spaces are possible by introducing mixed  $\bar{p}$ -quasi-norms; i.e., for  $\bar{p} = (p_1, p_2)$ ,

$$\|f\|_{\bar{p}} := \left( \int_{I_y} \left( \int_{I_x} |f(x, y)|^{p_1} dx \right)^{p_2/p_1} dy \right)^{1/p_2}.$$

The results with respect to those spaces will be reported elsewhere.

The paper is organized as follows: First we introduce anisotropic Besov spaces via directional moduli of smoothness and relate those spaces to anisotropic Besov spaces defined by averaged moduli and differences. In Section 3 we introduce two-parameter families of finite-dimensional subspaces and prove Jackson- and Bernstein estimates. Those estimates are basic for the following first characterization of anisotropic Besov spaces by one-parameter anisotropy-adapted linear approximation spaces in Section 4. Those characterizations are finally utilized to prove a general wavelet characterization result in Section 6. To this end we consider embedding and interpolation results for anisotropic Besov spaces in Section 5.

Generally, we use the notation  $a \lesssim b$  to abbreviate  $a \leq Cb$  with some constant  $C > 0$ ,  $a \gtrsim b$  if  $b \lesssim a$ , and for  $a \lesssim b \lesssim a$  we write  $a \sim b$ .

## 2. ANISOTROPIC BESOV SPACES

There are various possibilities for introducing isotropic and anisotropic Besov spaces. The most common ones are via the Fourier transform, via differences, and via moduli of smoothness. It is well-known for isotropic Besov spaces related to quasi-Banach spaces  $L_p$  with  $0 < p < 1$  if that those definitions do not always give the same spaces with equivalent quasi-norms; see [34]. Since anisotropic Besov spaces include in particular isotropic Besov spaces it is clear that this is also true for those. The most appropriate approach to deriving

wavelet characterizations is to study linear approximation spaces and to deduce Jackson and Bernstein estimates. Within our framework those estimates are essentially based on local error estimates for the approximation with respect to anisotropic polynomials, which involve certain moduli of smoothness. Thus it is natural to start with a definition of anisotropic Besov spaces via those moduli of smoothness. Contrary to the isotropic case we have to consider directional moduli of smoothness. We shall relate our definition for anisotropic Besov spaces to that via differences used in [29]; i.e., we shall argue that at least for the crucial range of parameters both definitions lead to the same function spaces with equivalent quasi-norms, which at least in the case of  $p \geq 1$  is well-known; see [26]. Hence, our final characterization results apply also to the anisotropic Besov spaces investigated in [28, 29].

### 2.1. Directional Moduli of Smoothness

Let  $\Omega$  denote the cube  $[0, 1]^2$ . For introducing anisotropic Besov spaces by directional moduli of smoothness we define at first inductively the directional partial difference operators  $\Delta_{h,i}^m$ ,  $m \in \mathbb{N}$ ,  $h \in \mathbb{R}$ ,  $i = 1, 2$ , by

$$(\Delta_{h,i}^1 f)(x) := f(x + he_i) - f(x), \quad \Delta_{h,i}^m := \Delta_{h,i}^1 \Delta_{h,i}^{m-1}, \quad m \geq 2,$$

where  $e_i$ ,  $i = 1, 2$ , denote the canonical basis vectors in  $\mathbb{R}^2$ . The directional differences are not defined for all  $x \in \Omega$  but are for  $x$  lying in the rectangles  $\Omega_{h,i}^m \subset \Omega$  with

$$\Omega_{h,i}^m := \{x \in \Omega \mid x + mhe_i \in \Omega\}.$$

For  $t \in \mathbb{R}^+$ ,  $m_i \in \mathbb{N}$ ,  $i = 1, 2$ , and  $0 < p \leq \infty$  the directional moduli of smoothness are given by

$$\omega_{m_i,i}(f, t, \Omega)_p := \sup_{|h| \leq t} \|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}, \quad f \in L_p(\Omega).$$

Anisotropic Besov spaces are then defined as follows.

**DEFINITION 2.1.** For  $\bar{a} = (a_1, a_2)$  with  $0 < a_2 \leq a_1 < 2$  and  $a_1 + a_2 = 2$ ,  $\bar{s} = (s_1, s_2)$  with  $s_1 := s/a_1$ ,  $s_2 := s/a_2$ , and  $s > 0$ , and  $(m_1, m_2) \in \mathbb{N}^2$  with  $m_1 > s_1$  and  $m_2 > s_2$ , we set for  $0 < p, q \leq \infty$

$$B_{p,q}^{\bar{s}} := B_q^{\bar{s}}(L_p(\Omega)) := \{f \in L_p(\Omega) \mid |f|_{B_{p,q}^{\bar{s}}} < \infty\},$$

where

$$|f|_{B_{p,q}^{\bar{s}}} := \left( \int_0^1 \left( \sum_{i=1}^2 t^{-s_i} \omega_{m_i,i}(f, t, \Omega)_p \right)^q \frac{dt}{t} \right)^{1/q}.$$

In addition to the (quasi-)seminorm  $|\cdot|_{B_{p,q}^{\bar{s}}}$  there is the (quasi-)norm

$$\|f\|_{B_{p,q}^{\bar{s}}} := \|f\|_p + |f|_{B_{p,q}^{\bar{s}}}.$$

## 2.2. Averaged Directional Moduli of Smoothness

In the context of isotropic Besov spaces it is known that the averaged moduli of smoothness are very helpful tools. In particular, they possess a summing property which resolves technical problems that may appear. Since similar problems arise in the anisotropic case we introduce analogously to the isotropic setting directional averaged moduli of smoothness: For  $t \in \mathbb{R}^+$ ,  $m_i \in \mathbb{N}$ ,  $i = 1, 2$ , we set

$$\tilde{\omega}_{m_i,i}(f, t, \Omega)_p := t^{-1} \int_{-t}^t \|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} dh, \quad f \in L_p(\Omega), \quad (1)$$

if  $1 \leq p < \infty$ , and

$$\tilde{\omega}_{m_i,i}(f, t, \Omega)_p := \left( t^{-1} \int_{-t}^t \|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}^p dh \right)^{1/p}, \quad f \in L_p(\Omega), \quad (2)$$

if  $0 < p < 1$ . Then the summing property reads as follows: For subdomains  $\Omega_j \subset \Omega$ ,  $j \in \{1, \dots, N\}$ ,  $N \in \mathbb{N}$ , with  $\cup_{j=1}^N \Omega_j = \Omega$  and a locally (with respect to  $N$  a uniformly) finite overlap, one has

$$\sum_{j=1}^N \tilde{\omega}_{m_i,i}(f, t, \Omega_j)_p^p \lesssim \tilde{\omega}_{m_i,i}(f, t, \Omega)_p^p, \quad f \in L_p(\Omega).$$

It turns out that both moduli of smoothness are equivalent for  $0 < p < \infty$ .

**THEOREM 2.1.** *Let  $f \in L_p(\Omega)$ ,  $0 < p < \infty$ . Then we have for  $0 < t \leq t_i := 1/4m_i$ ,  $i = 1, 2$ ,*

$$\omega_{m_i,i}(f, t, \Omega)_p \sim \tilde{\omega}_{m_i,i}(f, t, \Omega)_p. \quad (3)$$

*Proof.* We give only the sketch of a proof for  $1 \leq p < \infty$ . The case  $0 < p < 1$  can be treated similarly. First we note that the inequality “ $\gtrsim$ ” follows directly from

$$\|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} \leq \omega_{m_i,i}(f, t, \Omega)_p, \quad |h| \leq t.$$

Second, we observe that one has, analogously to an identity in [11, p. 184],

$$\Delta_{h,i}^{m_i} f(x) = \sum_{k=1}^{m_i} (-1)^k \binom{m_i}{k} [\Delta_{ks,i}^{m_i} f(x + k h e_i) - \Delta_{h+ks,i}^{m_i} f(x)],$$

which holds for  $s \in \mathbb{R}$  and for functions  $f$  defined on  $\mathbb{R}^2$ . Next one argues for each  $x_j$ ,  $j \neq i$ , as in [11, Chap. 6, Lemma 5.1], and integrates with respect to  $x_j$ , which gives for  $0 \leq \alpha_i < \beta_i \leq 1$ ,  $\delta_i := \beta_i - m_i \alpha_i - (2m_i)^{-1} > 0$ , and  $0 < h \leq t/2$ ,  $0 < t \leq 1/4m^2$ ,

$$\|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} \leq \frac{C}{t} \int_{\alpha_i m_i t}^{\beta_i m_i t} \|\Delta_{u,i}^{m_i} f\|_{p, \Omega_{u,i}^{m_i}} du, \quad f \in L_p(\Omega), \quad (4)$$

with  $C = C(m_i, \delta_i)$ . Thus, for  $|h| \leq t/2$  with  $0 < t \leq t_i/m_i$  it follows that

$$\|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} \leq \frac{C}{t} \int_{-m_i t}^{m_i t} \|\Delta_{u,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} du.$$

That is, one has

$$\omega_{m_i,i}(f, t, \Omega)_p \lesssim \tilde{\omega}_{m_i,i}(f, m_i t, \Omega)_p$$

and consequently

$$\omega_{m_i,i}(f, m_i t, \Omega)_p \lesssim \tilde{\omega}_{m_i,i}(f, m_i t, \Omega)_p$$

for  $0 < t \leq t_i/m_i$ . ■

Let us comment briefly on another observation with respect to the averaged directional moduli: They may be used to show that an integration of one-dimensional moduli gives terms which are equivalent to the directional moduli. This observation allows us, e.g., to transfer certain one-dimensional results directly to the anisotropic situation. As an application, in [19] we proved Marchaud-type inequalities for the directional moduli.

### 2.3. An Equivalence Result

Next we show the identity of anisotropic Besov spaces introduced via directional moduli of smoothness, averaged directional moduli, and directional differences.

**THEOREM 2.2.** *For  $\bar{a} = (a_1, a_2)$  with  $0 < a_2 \leq a_1 < \infty$ ,  $a_1 + a_2 = 2$ ,  $s_1 := s/a_1$ ,  $s_2 := s/a_2$ ,  $\bar{s} = (s_1, s_2)$  for  $s > 0$ , and  $0 < p < \infty$ ,  $0 < q \leq \infty$ , one has*

$$\begin{aligned} |f|_{B_{p,q}^{\bar{s}}(\Omega)} &\sim \|2^{s_1 j} \omega_{m_1,1}(f, 2^{-j}, \Omega)_p\|_{\ell_q} + \|2^{s_2 j} \omega_{m_2,2}(f, 2^{-j}, \Omega)_p\|_{\ell_q} \\ &\sim \|2^{s_1 j} \tilde{\omega}_{m_1,1}(f, 2^{-j}, \Omega)_p\|_{\ell_q} + \|2^{s_2 j} \tilde{\omega}_{m_2,2}(f, 2^{-j}, \Omega)_p\|_{\ell_q} \\ &\sim \left( \int_{-\infty}^{\infty} [|h|^{-s_1} \|\Delta_{h,1}^{m_1} f\|_{p, \Omega_{h,1}^{m_1}} + |h|^{-s_2} \|\Delta_{h,2}^{m_2} f\|_{p, \Omega_{h,2}^{m_2}}]^q \frac{dh}{h} \right)^{1/q}. \end{aligned} \quad (5)$$

*Proof.* The first two equivalences follow exactly as in the isotropic case by the monotonicity of the moduli; see, e.g., [11]. In the third equivalence the estimate “ $\gtrsim$ ” is rather obvious, because the difference terms are dominated by the moduli. Thus it remains to consider “ $\lesssim$ .” Again, we sketch the proof only for  $1 \leq p < \infty$ : Set  $\beta_i = 1$ ,  $\alpha_i m_i = 1/4$  such that  $\delta_i = \beta_i - m_i \alpha_i - (2m_i)^{-1} > 0$ . Then for  $0 < h \leq \tau = t/2$  and  $0 < t \leq t_i/2$  the inequality (4) gives

$$\|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} \lesssim \frac{1}{\tau} \int_{\tau/2}^{2m_i \tau} \|\Delta_{u,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}} du, \quad f \in L_p(\Omega).$$

Since  $\bar{\Omega}$  is compact we obtain

$$\begin{aligned} \int_0^\infty t^{-s_i q} \sup_{|h| \leq t} \|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}^q \frac{dt}{t} &\lesssim \int_0^{t_i/2} t^{-s_i q} \sup_{|h| \leq t} \|\Delta_{h,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}^q \frac{dt}{t} \\ &\lesssim \int_0^{t_i/2} t^{-s_i q - 1} \left( \int_{t/2}^{2m_i t} \|\Delta_{u,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}^q du \right) \frac{dt}{t} \\ &\lesssim \int_0^\infty t^{-s_i q} \|\Delta_{t,i}^{m_i} f\|_{p, \Omega_{h,i}^{m_i}}^q \frac{dt}{t}, \end{aligned}$$

which provides the desired estimate. ■

### 3. LINEAR ANISOTROPIC APPROXIMATION SPACES

The starting point is two possibly different one-dimensional biorthogonal multiresolutions in  $L_2(0, 1)$ , as described in, e.g., [5, 7]. In particular, we assume that we have so-called (primal and dual) scaling functions  $\varphi_{\lambda_i}^{(i)}: [0, 1] \rightarrow \mathbb{R}$  and  $\tilde{\varphi}_{\lambda_i}^{(i)}: [0, 1] \rightarrow \mathbb{R}$ ,  $i = 1, 2$ , with respect to nested sets of indices  $\lambda_i \in \Delta_j^{(i)}$ ,  $j \in \mathbb{N}_0$ , with  $\#\Delta_j^{(i)} \sim 2^j$ , which satisfy within  $L_2(0, 1)$  the biorthogonality relation

$$\langle \varphi_{\lambda_i}^{(i)}, \tilde{\varphi}_{\mu_i}^{(i)} \rangle = \delta_{\lambda_i, \mu_i}, \quad \lambda_i, \mu_i \in \Delta_j^{(i)}.$$

The supports of the scaling functions are supposed to be local; i.e., one has

$$|\text{supp } \varphi_{\lambda_i}^{(i)}| \sim |\text{supp } \tilde{\varphi}_{\lambda_i}^{(i)}| \sim 2^{-j}, \quad \lambda_i \in \Delta_j^{(i)}. \quad (6)$$

Moreover, we assume that for each  $\lambda_i \in \Delta_j^{(i)}$  there exist only finitely many  $\mu_i \in \Delta_j^{(i)}$ , such that the supports of  $\varphi_{\lambda_i}^{(i)}$  and  $\varphi_{\mu_i}^{(i)}$  and also those of  $\tilde{\varphi}_{\lambda_i}^{(i)}$  and  $\tilde{\varphi}_{\mu_i}^{(i)}$  overlap, and that there is an upper bound for this number independent of  $\lambda_i$  and  $j$ . Introducing for  $i \in \{1, 2\}$  and  $j \in \mathbb{N}_0$  the finite-dimensional subspaces  $S_j^{(i)} := \text{span}\{\varphi_{\lambda_i}^{(i)} \mid \lambda_i \in \Delta_j^{(i)}\}$  and  $\tilde{S}_j^{(i)} := \text{span}\{\tilde{\varphi}_{\lambda_i}^{(i)} \mid \lambda_i \in \Delta_j^{(i)}\}$ , one has  $S_j^{(i)} \subset S_{j+1}^{(i)}$  and  $\tilde{S}_j^{(i)} \subset \tilde{S}_{j+1}^{(i)}$ , respectively.

Furthermore we suppose that there are biorthogonal wavelet functions  $\psi_{\lambda_i}^{(i)}$  and  $\tilde{\psi}_{\lambda_i}^{(i)}$ ,  $\lambda_i \in \nabla_{j+1}^{(i)} := \Delta_{j+1}^{(i)} \setminus \Delta_j^{(i)}$ ,  $j \in \mathbb{N}_0$ , which are besides others basis functions for related finite-dimensional complements of  $S_j^{(i)}$  in  $S_{j+1}^{(i)}$  and of  $\tilde{S}_j^{(i)}$  in  $\tilde{S}_{j+1}^{(i)}$ , respectively; i.e., with  $W_{j+1}^{(i)} := \text{span}\{\psi_{\lambda_i}^{(i)} \mid \lambda_i \in \nabla_{j+1}^{(i)}\}$  and  $\tilde{W}_{j+1}^{(i)} := \text{span}\{\tilde{\psi}_{\lambda_i}^{(i)} \mid \lambda_i \in \nabla_{j+1}^{(i)}\}$  hold  $S_{j+1}^{(i)} = S_j^{(i)} \oplus W_{j+1}^{(i)}$ ,  $\tilde{S}_{j+1}^{(i)} = \tilde{S}_j^{(i)} \oplus \tilde{W}_{j+1}^{(i)}$ ,  $S_j^{(i)} \perp \tilde{W}_{j+1}^{(i)}$ , and  $\tilde{S}_j^{(i)} \perp W_{j+1}^{(i)}$ . Simplifying these notions we set  $\psi_{\lambda_i}^{(i)} := \varphi_{\lambda_i}^{(i)}$  and  $\tilde{\psi}_{\lambda_i}^{(i)} := \tilde{\varphi}_{\lambda_i}^{(i)}$  for  $\lambda_i \in \nabla_0^{(i)} := \Delta_0^{(i)}$ . The wavelet functions are supposed to possess local supports as in (6). We assume generally that all primal and dual scaling and wavelet functions are  $L_2(\Omega)$ -normalized and that their  $L_\infty(\Omega)$ -norms exist and are uniformly bounded by  $c2^{j/2}$  if  $\lambda_i \in \Delta_j^{(i)}$  and  $c$  is a fixed positive constant. Thus the wavelet functions permit in particular the so-called FWT, i.e., the fast wavelet transform. This means not only that the subspaces  $S_j^{(i)}$  and  $\tilde{S}_j^{(i)}$  are nested, but that each (single and dual) wavelet function related to some level can be represented by a finite linear combination of a number of scaling functions from the corresponding level such that these magnitudes are uniformly bounded, and the coefficients themselves are also uniformly bounded; e.g., we have for  $\lambda_i \in \nabla_{j+1}^{(i)}$ ,  $j \in \mathbb{N}_0$ , that  $\#\{\mu_i \in \Delta_{j+1}^{(i)} \mid \langle \psi_{\lambda_i}^{(i)}, \tilde{\varphi}_{\mu_i}^{(i)} \rangle \neq 0\} \lesssim 1$  and  $|\langle \psi_{\lambda_i}^{(i)}, \tilde{\varphi}_{\mu_i}^{(i)} \rangle| \lesssim 1$ . Ideally there is a representation formula which is independent of the level. Typically this is true for the (with respect to the interval  $[0, 1]$ ) inner scaling and inner wavelet functions, if they descend directly from a translation invariant setting and satisfy fixed refinement relations. Let us finally remark that our assumptions are general in the sense that all reasonable biorthogonal wavelets on an interval, which are constructed for applications in partial differential equations, fulfill them.

Now we move to the two-dimensional situation: For  $J = (j_1, j_2) \in \mathbb{N}_0^2$  we set  $\Delta_J = \Delta_{j_1}^{(1)} \times \Delta_{j_2}^{(2)}$  and for  $\bar{\lambda} = (\lambda_1, \lambda_2) \in \Delta_J$  we write  $\varphi_{\bar{\lambda}} := \varphi_{\lambda_1}^{(1)} \otimes \varphi_{\lambda_2}^{(2)}$ ,  $\tilde{\varphi}_{\bar{\lambda}} := \tilde{\varphi}_{\lambda_1}^{(1)} \otimes \tilde{\varphi}_{\lambda_2}^{(2)}$ ; i.e.,  $\varphi_{\bar{\lambda}}(x_1, x_2) = \varphi_{\lambda_1}^{(1)}(x_1)\varphi_{\lambda_2}^{(2)}(x_2)$  and  $\tilde{\varphi}_{\bar{\lambda}}(x_1, x_2) = \tilde{\varphi}_{\lambda_1}^{(1)}(x_1)\tilde{\varphi}_{\lambda_2}^{(2)}(x_2)$ , which are biorthogonal

with respect to the  $L_2(\Omega)$ -inner product. Then the finite-dimensional subspaces  $S_J := \text{span}\{\varphi_{\tilde{\lambda}} \mid \tilde{\lambda} \in \Delta_J\}$  can easily be identified with the algebraic tensor products  $S_{j_1}^{(1)} \otimes S_{j_2}^{(2)}$ .

Additionally, we introduce the following notation. By  $\mathcal{D}_J$ ,  $J = (j_1, j_2) \in \mathbb{N}_0^2$ , we denote the set of dyadic rectangles in  $\Omega$ ; i.e.,  $\mathcal{D}_J = \{2^{-j_1}[k_1, k_1 + 1] \times 2^{-j_2}[k_2, k_2 + 1], k_1 = 0, 1, \dots, 2^{j_1} - 1, k_2 = 0, 1, \dots, 2^{j_2} - 1\}$ , and we set  $\mathcal{D} := \bigcup_{J \in \mathbb{N}_0^2} \mathcal{D}_J$ . For each dyadic rectangle  $\square \in \mathcal{D}_J$  we set  $\ell_i(\square) := 2^{-j_i}$  and

$$\Delta_J(\square) := \{\tilde{\lambda} \in \Delta_J \mid |\square \cap \text{supp } \varphi_{\tilde{\lambda}}| > 0\},$$

as well as

$$\tilde{\square} := \bigcup_{\substack{\square \in \mathcal{D}_J \\ |\tilde{\square} \cap (\bigcup_{\tilde{\lambda} \in \Delta_J(\square)} \text{supp } \varphi_{\tilde{\lambda}})| > 0}} \square.$$

Then, the numbers of indices in  $\Delta_J(\square)$ ,  $\square \in \mathcal{D}_J$ , are uniformly bounded with respect to  $J \in \mathbb{N}_0^2$ ; i.e.,

$$\#\Delta_J(\square) \lesssim 1. \quad (7)$$

Furthermore, the number of dyadic subrectangles from  $\mathcal{D}_J$  building  $\tilde{\square}$  is uniformly bounded with respect to  $\square \in \mathcal{D}_J$  and  $J \in \mathbb{N}_0^2$ . In particular, our assumptions imply for  $\square \in \Delta_J$ ,  $J = (j_1, j_2)$ , that  $|\tilde{\square}| \sim 2^{-(j_1+j_2)}$  as well as  $\ell_i(\tilde{\square}) \sim 2^{-j_i}$ .

Within our context of anisotropic Besov spaces it is appropriate to introduce on  $\Omega := [0, 1]^2$  for  $M = (m_1, m_2) \in \mathbb{N}^2$ ,

$$\Pi^M := \left\{ P \in \Pi(\Omega) \mid P(x_1, x_2) = \sum_{k_1=0}^{m_1-1} \sum_{k_2=0}^{m_2-1} a_{k_1, k_2} x_1^{k_1} x_2^{k_2}, a_{k_1, k_2} \in \mathbb{R} \right\}.$$

That is,  $P \in \Pi^M$  is for each  $x_1 \in [0, 1]$  a polynomial in  $x_2$  of degree less than  $m_2$  and for each  $x_2 \in [0, 1]$  a polynomial in  $x_1$  of degree less than  $m_1$ . In the following, we shall assume that the subspaces  $S_J = \text{span}\{\varphi_{\tilde{\lambda}} \mid \tilde{\lambda} \in \Delta_J\}$  are *exact of order*  $M = (m_1, m_2)$ ; i.e.,  $\Pi^M \subset S_J$ . Generally, we assume for simplicity and for having spline wavelets “in the back of the head” that the functions in  $S_J$  are “piecewise” polynomials with respect to dyadic rectangles with sides of length  $2^{-j_1}$  and  $2^{-j_2}$ , respectively. We remark that the latter assumption is not necessary for most parts of the subsequent considerations. A generalization will be presented in a forthcoming paper.

After introducing all these notions and assumptions we proceed by proving Jackson and Bernstein estimates for the two-parameter family of finite-dimensional subspaces  $S_J$ . Beforehand we remark that one could give partly different proofs for  $1 \leq p < \infty$  and  $0 < p < 1$ ; see [19]. Proving results for  $0 < p < 1$  requires in most cases another proof as for  $1 \leq p < \infty$ , which is often caused by  $(L_p(\Omega))' = \{0\}$ , because this fact implies that duality arguments cannot be used directly. On the other hands proofs for  $0 < p < 1$  often apply with only slight modifications to  $1 \leq p < \infty$  also. For example, the standard proof for the approximation properties of polynomial subspaces of anisotropic type in the case  $1 \leq p < \infty$  cannot be transferred to the case  $0 < p < 1$ , but a slight modification of our proof for  $0 < p < 1$  can be applied to the other case also.

Another difficulty in the case  $0 < p < 1$  is that the projectors  $P_J$ ,  $J = (j_1, j_2) \in \mathbb{N}_0^2$ , defined by

$$P_J f := \sum_{\tilde{\lambda} \in \Delta_J} \langle f, \tilde{\varphi}_{\tilde{\lambda}} \rangle \varphi_{\tilde{\lambda}}, \quad f \in L_p(\Omega), \quad (8)$$

are not uniformly bounded as operators from  $L_p(\Omega)$  to  $L_p(\Omega)$ .

Contrarily, since the projection operators  $P_J$  on  $S_J$  turn out to be bounded for  $p \geq 1$  the following approach can be applied in that case: Starting with an approximation result for anisotropic polynomials with respect to some appropriate moduli, see, e.g., [3, 8] or (9), by a Bernstein estimate one obtains a characterization of anisotropic Besov spaces via linear approximation spaces. Then the boundedness of projection operators and a representation of differences between those operators in terms of suitable wavelets lead directly to the goal, i.e., to a wavelet characterization. Therefore, to shorten our presentation, we present in the following complete proofs only for the case  $0 < p < 1$ .

### 3.1. Jackson Estimates

Given  $f \in L_p(\Omega)$ ,  $0 < p < \infty$ , the approximation error with respect to  $S_J$  will be denoted by

$$E_J(f)_p := \inf_{\chi \in S_J} \|f - \chi\|_p.$$

Jackson estimates for the approximation error are typically the result of the fact that function systems with well-known approximation properties are locally reproduced by the scaling functions.

Related to the anisotropic order  $M$  of exactness of  $S_J$ , the first step is to generalize the well-known estimate

$$\inf_{P \in \Pi^M} \|f - P\|_{p,\Omega} \lesssim \sum_{i=1,2} \omega_{m_i,i}(f, 1, \Omega)_p, \quad f \in L_p(\Omega), p \geq 1 \quad (9)$$

(cf. [3, 8]), to the case  $0 < p < 1$ , which requires further notions. In particular, we need the so-called mixed differences and the corresponding mixed moduli of smoothness: For  $\bar{h} = (h_1, h_2) \in \mathbb{R}_+^2$  and  $\bar{m} = (m_1, m_2) \in \mathbb{N}^2$  we write

$$\Delta_{\bar{h}}^{\bar{m}} f(x) = \Delta_{h_2,2}^{m_2} \circ \Delta_{h_1,1}^{m_1} f(x),$$

and for  $\bar{t} = (t_1, t_2) \in \mathbb{R}_+^2$  we set

$$\omega_{\bar{m}}(f, \bar{t}, \Omega)_p := \sup_{\bar{h} \leq \bar{t}} \|\Delta_{\bar{h}}^{\bar{m}} f\|_{p, \Omega_{\bar{h}}^{\bar{m}}},$$

where

$$\Omega_{\bar{h}}^{\bar{m}} := \{x = (x_1, x_2) \in \Omega \mid x + m_1 h_1 e_1 + m_2 h_2 e_2 \in \Omega\}.$$

LEMMA 3.1. For  $\bar{n} \leq \bar{m}$  it holds that

$$\omega_{\bar{m}}(f, \bar{t}, \Omega)_p \leq 2^{\sum_{i=1}^2 m_i - n_i} \omega_{\bar{n}}(f, \bar{t}, \Omega)_p, \quad (10)$$



if  $1 \leq p \leq \infty$ , and

$$\omega_{\bar{m}}(f, \bar{t}, \Omega)_p^p \leq 2^{\sum_{i=1}^2 m_i - n_i} \omega_{\bar{n}}(f, \bar{t}, \Omega)_p^p, \quad (11)$$

if  $0 < p < 1$ .

*Proof.* Setting  $T_{h,i} f(x) := f(x + h e_i)$  one has  $\int_G |T_{h,i} g(x)|^p dx = \int_{G+h e_i} |g(x)|^p dx$ . For  $0 < p < 1$  it holds that

$$\|\Delta_{\bar{h}}^{\bar{m}} f\|_{p, \Omega_{\bar{h}}^{\bar{m}}}^p = \int_{\Omega_{\bar{h}}^{\bar{m}}} |(I - T_{h_2, 2})^{m_2} (I - T_{h_1, 1})^{m_1} f(x)|^p dx \leq 2^{\sum_{i=1}^2 m_i - n_i} \|\Delta_{\bar{h}}^{\bar{n}} f\|_{p, \Omega_{\bar{h}}^{\bar{n}}}^p,$$

which implies (11) by

$$\begin{aligned} \omega_{\bar{m}}(f, \bar{t}, \Omega)_p^p &= \sup_{\bar{h} \leq \bar{i}} \|\Delta_{\bar{h}}^{\bar{m}} f\|_{p, \Omega_{\bar{h}}^{\bar{m}}}^p \leq 2^{\sum_{i=1}^2 m_i - n_i} \sup_{\bar{h} \leq \bar{i}} \|\Delta_{\bar{h}}^{\bar{n}} f\|_{p, \Omega_{\bar{h}}^{\bar{n}}}^p \\ &= 2^{\sum_{i=1}^2 m_i - n_i} \omega_{\bar{n}}(f, \bar{t}, \Omega)_p^p. \end{aligned}$$

If  $1 \leq p \leq \infty$  one gets by the triangle inequality that

$$\|\Delta_{\bar{h}}^{\bar{m}} f\|_{p, \Omega_{\bar{h}}^{\bar{m}}} \leq 2^{\sum_{i=1}^2 m_i - n_i} \|\Delta_{\bar{h}}^{\bar{n}} f\|_{p, \Omega_{\bar{h}}^{\bar{n}}},$$

which gives (10). ■

Next we make essential use of the following result from [31]: For  $G = [a_1, b_1] \times [a_2, b_2]$  and  $f \in L_p(G)$  there is a constant  $C = C(f, G)$  such that for any  $k \in \mathbb{N}$  there are  $c_k \in \mathbb{R}$  with

$$\begin{aligned} \|f - C\|_{p, G}^p &\leq c_k \left\{ h_1^{-1} \int_0^{h_1} \|\Delta_{t, 1}^k f\|_{p, G_{t, 1}^k}^p dt + h_2^{-1} \int_0^{h_2} \|\Delta_{t, 2}^k f\|_{p, G_{t, 2}^k}^p dt \right. \\ &\quad \left. + \|\Delta_{h_1, 1}^1 f\|_{p, G_{(h_1, 1)}^{(1, 1)}}^p + \|\Delta_{h_2, 2}^1 f\|_{p, G_{(h_1, h_2)}^{(1, 1)}}^p + \|\Delta_{h_1, 1}^{(1, 1)} f\|_{p, G_{(h_1, h_2)}^{(1, 1)}}^p \right\}, \end{aligned} \quad (12)$$

where  $h_i = (b_i - a_i)/k$ ,  $i = 1, 2$ .

**THEOREM 3.1.** For  $0 < p < 1$

$$\inf_{P \in \Pi^M} \|f - P\|_{p, \Omega} \lesssim \sum_{i=1}^2 \omega_{m_i, i}(f, 1, \Omega)_p, \quad f \in L_p(\Omega), \quad (13)$$

holds.

*Proof.* We construct a polynomial  $P = P(f) = \sum_{k_1=0}^{m_1-1} \sum_{k_2=0}^{m_2-1} a_{k_1, k_2} x_1^{k_1} x_2^{k_2} \in \Pi^M$  such that

$$\|f - P\|_{p, \Omega} \lesssim \sum_{i=1}^2 \omega_{m_i, i}(f, 1, \Omega)_p.$$

To this end we set  $n := m_1 + m_2$ ,  $\delta := 1/2n$ , and  $\Omega_\ell := \Omega_{(\delta, \delta)}^{(\ell, \ell)}$  for  $\ell = 1, \dots, \max(m_1 - 1, m_2 - 1)$ . In a first step we fix the coefficients of the monomials of degree  $n - 2$  by defining

$$a_{m_1-1, m_2-1} = C(\Delta_{(\delta, \delta)}^{(m_1-1, m_2-1)} f, \Omega_{\max(m_1-1, m_2-1)}) \frac{\delta^{-n+2}}{(n-2)!}.$$

Next we set

$$f_1(x_1, x_2) := f(x_1, x_2) - a_{m_1-1, m_2-1} x_1^{m_1-1} x_2^{m_2-1}.$$

Generally we write, for  $1 \leq i \leq n-2$ ,

$$f_i(x_1, x_2) := f_{i-1}(x_1, x_2) - \sum_{j=\max(0, m_2-i)}^{\min(n-1-i, m_2-1)} a_{n-1-i-j, j} x_1^{n-1-i-j} x_2^j$$

and, for  $j = \max(0, m_2-i-1), \dots, \min(n-2-i, m_2-1)$ ,

$$a_{n-2-i, j} := C(\Delta_{(\delta, \delta)}^{(n-2-i-j, j)}) f_i, \Omega_{\max(n-2-i-j, j)} \frac{\delta^{-n+2+i}}{(n-2-i)!}.$$

The polynomial  $P \in \Pi^M$  is then defined by

$$P(x_1, x_2) := \sum_{i=1}^{n-1} \sum_{j=\max(0, m_2-i)}^{\min(n-1-i, m_2-1)} a_{n-1-i-j, j} x_1^{n-1-i-j} x_2^j.$$

With these definitions we obtain

$$\|f - P\|_{p, \Omega}^p = \|f_{n-2} - a_{0,0}\|_{p, \Omega}^p = \|f_{n-2} - C(f_{n-2}, \Omega)\|_{p, \Omega}^p.$$

Next we apply (12) with  $k = 2n$  and note that  $\Delta_{\delta,1}^{2n} f_{n-2} = \Delta_{\delta,1}^{2n} f$ ,  $\Delta_{\delta,2}^{2n} f_{n-2} = \Delta_{\delta,2}^{2n} f$ , and  $\Delta_{(\delta,\delta)}^{(1,1)} f_{n-2} = \Delta_{(\delta,\delta)}^{(1,1)} f_{n-3}$ . Therefore we obtain

$$\begin{aligned} \|f - P\|_{p, \Omega}^p &\lesssim \delta^{-1} \int_0^\delta \|\Delta_{t,1}^{2n} f\|_{p, \Omega_{t,1}^{2n}}^p dt + \delta^{-1} \int_0^\delta \|\Delta_{t,2}^{2n} f\|_{p, \Omega_{t,2}^{2n}}^p dt \\ &\quad + \|\Delta_{\delta,1}^1 f_{n-2}\|_{p, \Omega_1}^p + \|\Delta_{\delta,2}^1 f_{n-2}\|_{p, \Omega_1}^p + \|\Delta_{(\delta,\delta)}^{(1,1)} f_{n-3}\|_{p, \Omega_1}^p. \end{aligned}$$

Next we apply (12) with respect to  $k = 2n-1$  and obtain for the third term

$$\begin{aligned} \|\Delta_{\delta,1}^1 f_{n-2}\|_{p, \Omega_1}^p &= \|\Delta_{\delta,1}^1 f_{n-3} - a_{1,0} \delta\|_{p, \Omega_1}^p \\ &= \|\Delta_{\delta,1}^1 f_{n-3} - C(\Delta_{\delta,1}^1 f_{n-3}, \Omega_1)\|_{p, \Omega_1}^p \\ &\lesssim \delta^{-1} \int_0^\delta \|\Delta_{t,1}^{2n-1} f\|_{p, \Omega_{t,1}^{2n-1}}^p dt + \delta^{-1} \int_0^\delta \|\Delta_{t,2}^{2n-1} f\|_{p, \Omega_{t,2}^{2n-1}}^p dt \\ &\quad + \|\Delta_{\delta,1}^2 f_{n-3}\|_{p, \Omega_2}^p + \|\Delta_{(\delta,\delta)}^{(1,1)} f_{n-3}\|_{p, \Omega_2}^p + \|\Delta_{(\delta,\delta)}^{(2,1)} f_{n-4}\|_{p, \Omega_2}^p. \end{aligned}$$

The fourth term gives similarly

$$\begin{aligned} \|\Delta_{\delta,2}^1 f_{n-2}\|_{p, \Omega_1}^p &= \|\Delta_{\delta,2}^1 f_{n-3} - a_{0,1} \delta\|_{p, \Omega_1}^p \\ &= \|\Delta_{\delta,2}^1 f_{n-3} - C(\Delta_{\delta,2}^1 f_{n-3}, \Omega_1)\|_{p, \Omega_1}^p \\ &\lesssim \delta^{-1} \int_0^\delta \|\Delta_{t,2}^{2n-1} f\|_{p, \Omega_{t,2}^{2n-1}}^p dt + \delta^{-1} \int_0^\delta \|\Delta_{t,2}^{2n-1} f\|_{p, \Omega_{t,2}^{2n-1}}^p dt \\ &\quad + \|\Delta_{\delta,2}^2 f_{n-3}\|_{p, \Omega_2}^p + \|\Delta_{(\delta,\delta)}^{(1,1)} f_{n-3}\|_{p, \Omega_2}^p + \|\Delta_{(\delta,\delta)}^{(1,2)} f_{n-4}\|_{p, \Omega_2}^p. \end{aligned}$$

Repeating this procedure at most  $n$  times and replacing averaged differences with moduli lead finally by Lemma 3.1 to

$$\begin{aligned} \|f - P_{n-2}\|_{p,\Omega}^p &\lesssim \sum_{i=1}^2 \sum_{\ell=0}^{n-1} \omega_{2n-\ell,i}(f, 1, \Omega)_p^p \\ &\quad + \sum_{\ell=0}^{m_2} \|\Delta_{(\delta,\delta)}^{(m_1,\ell)} f\|_{\Omega_{\max(m_1,\ell)}}^p + \sum_{\ell=0}^{m_1-1} \|\Delta_{(\delta,\delta)}^{(\ell,m_1)} f\|_{\Omega_{m_2}}^p \\ &\lesssim \sum_{i=1}^2 \omega_{m_i,i}(f, 1, \Omega)_p^p. \blacksquare \end{aligned}$$

If  $\Omega_\rho$ ,  $\rho = (\rho_1, \rho_2) \in \mathbb{R}_+^2$ , denotes a rectangle with side lengths  $\rho_1$  and  $\rho_2$ , a scaling argument gives by (13) that

$$\inf_{P \in \Pi^M} \|f - P\|_{p,\Omega_\rho} \lesssim \sum_{i=1}^2 \omega_{m_i,i}(f, \rho_i, \Omega_\rho)_p, \quad f \in L_p(\Omega_\rho). \quad (14)$$

As a first illustrative application of Theorem 3.1 we give a generalization of a well-known isotropic result, which can be found in [11].

**THEOREM 3.2.** *Let  $f \in L_p(\Omega)$ ,  $0 < p < 1$ . If  $\omega_{m_1,1}(f, t, \Omega)_p = \omega_{m_2,2}(f, t, \Omega)_p = 0$  for some  $0 < t \leq 1$ , then  $f$  is on  $\Omega$  equivalent to a polynomial in  $\Pi^M$ ,  $M = (m_1, m_2)$ .*

*Proof.* First we note that  $\omega_{m_i,i}(f, t, \Omega)_p = 0$  for some  $t \in (0, 1]$  implies that  $\omega_{m_i,i}(f, \cdot, \Omega)_p = 0$  on  $(0, 1]$ . In fact,  $\omega_{m_1,1}(f, t, \Omega)_p = 0$  implies for a.e.  $x_2 \in I_2$  that  $\omega_{m_1}(f(\cdot, x_2), t, I_1)_p = 0$ . This gives by arguments in [11, p. 370]; for a.e.  $x_2 \in I_2$ , that  $\omega_{m_1}(f(\cdot, x_2), t, I_1)_p = 0$  for  $t \in (0, 1]$ , which enforces  $\omega_{m_1,1}(f, \cdot, \Omega)_p = 0$  identically on  $(0, 1]$ .

Clearly, the same arguments work for  $i = 2$ . Therefore, the proof of Theorem 3.1 provides a polynomial  $P \in \Pi^M$  with  $\|f - P\|_{p,\Omega} = 0$ ; i.e.,  $f \in \Pi^M$ .  $\blacksquare$

*Remark.* For continuous functions  $f$  the result of Theorem 3.2 can be verified without using Theorem 3.1: The assumption  $\omega_{m_1,1}(f, t, \Omega)_p = 0$  implies for  $\varphi \in C_0^\infty(\Omega)$  and  $h \in (0, 1)$  that

$$\int_{\Omega_{-h,1}^{m_1}} f(x) h^{-m_1} \Delta_{-h,1}^{m_1} \varphi(x) dx = \int_{\Omega_{h,1}^{m_1}} h^{-m_1} \Delta_{h,1}^{m_1} f(x) \varphi(x) = 0.$$

Then  $h \rightarrow 0+$  gives

$$\int_{\Omega} f(x) \frac{\partial^{m_1}}{\partial x_1^{m_1}} \varphi(x) dx = 0. \quad (15)$$

Analogously, it follows that

$$\int_{\Omega} f(x) \frac{\partial^{m_2}}{\partial x_2^{m_2}} \varphi(x) dx = 0. \quad (16)$$

The relation (15) implies that for a.e.  $x_2$  there is a polynomial  $p_{x_2} \in \Pi_{m_1-1}$  such that  $f(\cdot, x_2) = p_{x_2}$  on  $I$ , and the relation (16) gives for a.e.  $x_1$  a polynomial  $p_{x_1} \in \Pi^{m_2}$  such

that  $f(x_1, \cdot) = p_{x_1}$  on  $I$ . Thus the continuous representation of  $f$  is a.e. a polynomial in  $\Pi^{m_1, m_2}$ , hence is very everywhere: Let  $f(t, x_2) = \sum_{k=0}^{m_1-1} a_k(x_2)t^k$  and  $f(x_1, s) = \sum_{\ell=0}^{m_2-1} b_\ell(x_1)s^\ell$  with appropriate functions  $a_k$  and  $b_\ell$ . Then one has

$$\sum_{k=0}^{m_1-1} a_k(x_2)t^k = \sum_{\ell=0}^{m_2-1} b_\ell(x_1)s^\ell,$$

which leads for appropriate parameters to a Vandermondsche determinant and thus provides the assertion.

Next we denote by  $\Pi_J^M$  the finite-dimensional vector spaces of “piecewise” polynomials of degree less than  $M = (m_1, m_2)$  on  $\mathcal{D}_J$ ,  $J = (j_1, j_2)$ . Furthermore, we define for an arbitrary subdomain  $\Omega' \subset \Omega$  and for  $f \in L_p(\Omega)$

$$E_M(f, \Omega')_p := \inf_{P \in \Pi^M} \|f - P\|_{p, \Omega'}.$$

Since each  $S \in \Pi_J^M$  belongs to  $L_2(\Omega)$  we may set

$$P_J S = \sum_{\tilde{\lambda} \in \Delta_J} \langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle \varphi_{\tilde{\lambda}}.$$

Furthermore, we still assume that  $M = (m_1, m_2)$  denotes the anisotropic order of polynomial exactness in  $S_J$ .

LEMMA 3.2. *Let  $0 < p < 1$ . Then one has for  $\square \in \mathcal{D}_J$  that*

$$\|P_J S\|_{p, \square} \lesssim \|S\|_{p, \tilde{\square}}, \quad S \in \Pi_J^M, \quad (17)$$

and

$$\|S - P_J S\|_{p, \square} \lesssim E_M(S, \tilde{\square})_p, \quad S \in \Pi_J^M. \quad (18)$$

*Proof.* Suppose (17) is proved, then (18) follows by the fact that  $P_J(P) = P$  for all polynomials  $P \in \Pi^M$ . In fact, if  $P \in \Pi^M$  is a polynomial near-best  $L_p(\tilde{\square})$ -approximation to  $S$ , then

$$\begin{aligned} \|S - P_J(S)\|_{p, \square}^p &\lesssim \|S - P\|_{p, \square}^p + \|P - P_J(S)\|_{p, \square}^p \\ &\lesssim \|S - P\|_{p, \tilde{\square}}^p \\ &\lesssim E_M(S, \tilde{\square})_p^p. \end{aligned}$$

Therefore it remains to prove (17). To this end we take an arbitrary  $S \in \Pi_J^M$ . Then  $P_J S$  is defined and one has

$$P_J S|_{\square} = \sum_{\tilde{\lambda} \in \Delta_J(\square)} \langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle \varphi_{\tilde{\lambda}}|_{\square}.$$

The equivalence of quasi-norms on finite-dimensional vector spaces and the property  $|\tilde{\square}| \sim 2^{-(j_1+j_2)}$  for  $\square \in \Delta_J$  give a constant  $c = c(2, p)$  such that

$$\|S\|_{2,\tilde{\square}} \leq c 2^{(j_1+j_2)(1/p-1/2)} \|S\|_{p,\tilde{\square}}, \quad S \in \Pi_J^M,$$

which implies for  $\tilde{\lambda} \in \Delta_J(\square)$  that

$$|\langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle| \leq \|S\|_{2,\text{supp}(\tilde{\varphi}_{\tilde{\lambda}})} \leq \|S\|_{2,\tilde{\square}} \leq 2^{(j_1+j_2)(1/p-1/2)} \|S\|_{p,\tilde{\square}}. \quad (19)$$

Therefore, we obtain by  $\ell_p \hookrightarrow \ell_1$ ,  $0 < p \leq 1$ , and by Hölder's inequality that

$$\begin{aligned} \|P_J S\|_{p,\square}^p &= \int_{\square} \left| \sum_{\tilde{\lambda} \in \Delta_J(\square)} \langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle \varphi_{\tilde{\lambda}}(x) \right|^p dx \\ &\leq \sum_{\tilde{\lambda} \in \Delta_J(\square)} |\langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p \int_{\square} |\varphi_{\tilde{\lambda}}(x)|^p dx \\ &\lesssim 2^{-(j_1+j_2)(1-p/2)} \sum_{\tilde{\lambda} \in \Delta_J(\square)} |\langle S, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p \\ &\lesssim \|S\|_{p,\tilde{\square}}^p. \quad \blacksquare \end{aligned} \quad (20)$$

As an immediate consequence of Lemma 3.2 we get a stability property in  $S_J$ .

LEMMA 3.3. *One has uniformly in  $J \in \mathbb{N}^2$*

$$\|v\|_{p,\Omega} \sim \left( \sum_{\lambda \in \Delta_J} |\langle v, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p 2^{(j_1+j_2)(p/2-1)} \right)^{1/p}, \quad v \in S_J.$$

*Proof.* “ $\gtrsim$ ” Since  $S_J \subset \Pi_J^M$  we get for  $v \in S_J$  by (19) and (7) that

$$\sum_{\lambda \in \Delta_J} |\langle v, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p 2^{(j_1+j_2)(p/2-1)} \lesssim \sum_{\square \in \mathcal{D}_J} \|v\|_{p,\tilde{\square}}^p \lesssim \|v\|_{p,\Omega}^p.$$

“ $\lesssim$ ” Since  $P_J v = v$  for  $v \in S_J$  we get by (20) that

$$\|v\|_{p,\square}^p \lesssim 2^{(j_1+j_2)(p/2-1)} \sum_{\tilde{\lambda} \in \Delta_J(\square)} |\langle v, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p,$$

which gives the assertion by summing up with respect to  $\square$ .  $\blacksquare$

Next we prove the Jackson estimate for the approximation of  $f \in L_p(\Omega)$  with respect to  $S_J$ .

THEOREM 3.3. *One has for  $J = (j_1, j_2) \in \mathbb{N}_0^2$  that*

$$E_J(f)_p \lesssim \sum_{i=1}^2 \omega_{m_i,i}(f, 2^{-j_i}, \Omega), \quad f \in L_p(\Omega). \quad (21)$$

*Proof.* First we introduce an appropriate anisotropic approximation procedure for a given  $f \in L_p(\Omega)$ : For each  $\square \in \mathcal{D}_J$  and  $f \in L_p(\Omega)$  let  $P_\square$  be a near-best  $L_p(\square)$ -approximation to  $f$  from  $\Pi^M$ . If  $P \in \Pi^M$  is a near-best approximation of  $f$  on  $\tilde{\square}$  we get

$$\|P_\square - P\|_{p,\square} \lesssim \|P_\square - f\|_{p,\square} + \|f - P\|_{p,\square} \lesssim E_M(f, \tilde{\square})_p.$$

Next we observe that there is a constant  $c > 0$  depending on  $p$ ,  $M$ , and the constant in  $|\tilde{\square}| \lesssim 2^{-(j_1+j_2)}$  such that uniformly in  $\square \in \mathcal{D}$

$$\|P\|_{p,\tilde{\square}} \leq c \|P\|_{p,\square}, \quad P \in \Pi^M.$$

This fact follows simply from the equivalence of quasi-norms on  $\Pi^M$  and a change of variables. Consequently, we get

$$\|P_\square - P\|_{p,\tilde{\square}} \lesssim \|P_\square - P\|_{p,\square} \lesssim E_M(f, \tilde{\square})_p,$$

which gives

$$\|f - P_\square\|_{p,\tilde{\square}} \lesssim \|f - P\|_{p,\tilde{\square}} + \|P - P_\square\|_{p,\tilde{\square}} \lesssim E_M(f, \tilde{\square})_p.$$

Next we define  $S_J \in \Pi_J^M \in L_2(\Omega)$  by

$$S_J(x) := P_\square(x), \quad x \in \square, \quad \square \in \mathcal{D}_J,$$

and an approximation of  $f$  in  $S_J$  by

$$A_J(f) := P_J S_J.$$

This definition implies by (18) for  $\square \in \Delta_J$  that

$$\begin{aligned} \|f - A_J(f)\|_{p,\square} &\lesssim \|f - S_J\|_{p,\square} - \|S_J - P_J S_J\|_{p,\square} \\ &\lesssim \|f - P_\square\|_{p,\square} + E_M(S_J, \tilde{\square})_p \\ &\lesssim E_M(f, \tilde{\square})_p + E_M(S_J, \tilde{\square})_p. \end{aligned} \tag{22}$$

Since for any  $\square' \in \mathcal{D}_J$  with  $\square' \subset \tilde{\square}$  it holds that

$$\begin{aligned} \|S_J - P_\square\|_{p,\square'} &= \|P_{\square'} - P_\square\|_{p,\square'} \leq \|f - P_{\square'}\|_{p,\square'} + \|f - P_\square\|_{p,\square'} \\ &\lesssim E_M(f, \square')_p + \|f - P_\square\|_{p,\tilde{\square}} \lesssim E_M(f, \tilde{\square})_p, \end{aligned}$$

we obtain

$$\begin{aligned} E_M(S_J, \tilde{\square})_p^p &\leq \|S_J - P_\square\|_{p,\tilde{\square}}^p = \sum_{\square' \subset \tilde{\square}} \|S_J - P_\square\|_{p,\square'}^p \\ &\lesssim \sum_{\square' \subset \tilde{\square}} E_M(f, \square')_p^p \lesssim E_M(f, \tilde{\square})_p^p, \end{aligned}$$

because the number of  $\square' \in \mathcal{D}_J$  with  $\square' \subset \tilde{\square}$ ,  $\square \in \mathcal{D}_J$ , is uniformly bounded. Thus, (22) implies

$$\|f - A_J(f)\|_{p,\square} \lesssim E_M(f, \tilde{\square})_p.$$

Finally, we apply Theorem 3.1, which gives

$$\begin{aligned} \|f - A_J(f)\|_{p,\Omega}^p &= \sum_{\square \in \mathcal{D}_J} \|f - A_J(f)\|_{p,\square}^p \lesssim \sum_{\square \in \mathcal{D}_J} E_M(f, \tilde{\square})_p^p \\ &\lesssim \sum_{i=1}^2 \sum_{\square \in \mathcal{D}_J} \omega_{m_i,i}(f, \ell_i(\tilde{\square}), \tilde{\square})_p^p \lesssim \sum_{i=1}^2 \omega_{m_i,i}(f, 2^{-j_i}, \Omega)_p^p, \end{aligned} \quad (23)$$

where we again repeatedly apply the equivalence of the moduli of smoothness to the averaged moduli, the summing property of the averaged moduli, and  $\ell_i(\tilde{\square}) \lesssim 2^{-j_i}$ . ■

From Theorem 3.3 follows directly the Jackson estimate for  $f \in B_{p,q}^{\bar{s}}$ ,  $0 < p < 1$ ,  $0 < q < \infty$ :

**THEOREM 3.4.** *Under the above assumptions on the scaling function, it holds for  $0 < \bar{s} < M$  and  $0 < p < 1$ ,  $0 < q < \infty$ , that*

$$\|f - A_J f\|_p \lesssim (2^{-j_1 s_2} + 2^{-j_2 s_2}) |f|_{B_{p,q}^{\bar{s}}}, \quad f \in B_{p,q}^{\bar{s}}. \quad (24)$$

Since similar proofs give the same estimates for  $1 \leq p < \infty$  we end up with the following corollary.

**COROLLARY 3.1.** *For  $0 < p < \infty$  and  $J = (j_1, j_2) \in \mathbb{N}_0^2$  it holds that*

$$E_J(f)_p \lesssim \sum_{i=1}^2 \omega_{m_i,i}(f, 2^{-j_i}, \Omega), \quad f \in L_p(\Omega), \quad (25)$$

as well as

$$\|f - A_J f\|_p \lesssim (2^{-j_1 s_1} + 2^{-j_2 s_2}) |f|_{B_{p,q}^{\bar{s}}}, \quad f \in B_{p,q}^{\bar{s}}. \quad (26)$$

### 3.2. Bernstein Estimates

Bernstein estimates are typically proved under appropriate smoothness assumptions on the basis functions. It turns out that within our framework the following formulation is appropriate. We say that *the scaling functions are  $B_{p,q}^{\bar{s}}$ -smooth*,  $\bar{s} > 0$ , if for  $(r_1, r_2) > \bar{s}$  there are  $\tilde{\lambda}_J = (\lambda_{1,J}, \lambda_{2,J}) \in \Delta_J$  such that uniformly in  $J \in \mathbb{N}_0^2$  and  $k \in \mathbb{N}_0$ ,

$$\omega_{r_i,i}(\varphi_{\tilde{\lambda}_J}^-, 2^{-k}, \Omega)_p \lesssim \omega_{r_i,i}(\varphi_{\tilde{\lambda}_J}^-, 2^{-k}, \Omega)_p, \quad \tilde{\lambda}_J \in \Delta_J,$$

and if uniformly in  $J = (j_1, j_2) \in \mathbb{N}_0^2$  it holds that

$$|\varphi_{\tilde{\lambda}_J}^-|_{B_{p,q}^{\bar{s}}} \lesssim 2^{(j_1+j_2)(1/2-1/p)} \sum_{i=1}^2 2^{j_i s_i}. \quad (27)$$

Both assumptions are satisfied by tensor products of typical wavelet constructions; in particular, they are fulfilled within the translation-invariant context if the starting scaling functions are smooth enough. Let us emphasize that the wavelet-characterization result for the Besov spaces under consideration typically enforces no further approximation or smoothness assumptions on the dual scaling and wavelet functions. Consequently, we do not need all the properties of, e.g., the biorthogonal wavelet bases constructed in [9].

**THEOREM 3.5.** *If the scaling functions are  $B_{p,q}^{\bar{s}}$ -smooth for  $0 < p < \infty$  and  $0 < q \leq \infty$  then one has uniformly in  $J = (j_1, j_2) \in \mathbb{N}_0^2$  that*

$$|f|_{B_{p,q}^{\bar{s}}} \lesssim (2^{s_1 j_1} + 2^{s_2 j_2}) \|f\|_p, \quad f \in S_J. \quad (28)$$

*Proof.* We present only the proof for  $0 < p < 1$ , since the proof for  $1 \leq p < \infty$  needs only slight modifications. Let  $f = \sum_{\bar{\lambda} \in \Delta_J} a_{\bar{\lambda}} \varphi_{\bar{\lambda}} \in S_J$ . Then we obtain for  $r_i > s_i$  and  $k \in \mathbb{N}$  by standard estimates and the  $B_{p,q}^{\bar{s}}$ -smoothness of the scaling functions that

$$\begin{aligned} \omega_{r_i,i}(f, 2^{-k}, \Omega)_p &= \sup_{|h| \leq 2^{-k}} \left( \int_{\Omega_{h,i}^{r_i}} |\Delta_{h,i}^{r_i} f(x)|^p dx \right)^{1/p} \\ &= \sup_{|h| \leq 2^{-k}} \left( \int_{\Omega_{h,i}^{r_i}} \left| \Delta_{h,i}^{r_i} \sum_{\bar{\lambda} \in \Delta_J} a_{\bar{\lambda}} \varphi_{\bar{\lambda}}(x) \right|^p dx \right)^{1/p} \\ &\leq \sup_{|h| \leq 2^{-k}} \left( \sum_{\bar{\lambda} \in \Delta_J} |a_{\bar{\lambda}}|^p \int_{\Omega_{h,i}^{r_i}} |\Delta_{h,i}^{r_i} \varphi_{\bar{\lambda}}(x)|^p dx \right)^{1/p} \\ &\leq \left( \sum_{\bar{\lambda} \in \Delta_J} |a_{\bar{\lambda}}|^p \sup_{|h| \leq 2^{-k}} \int_{\Omega_{h,i}^{r_i}} |\Delta_{h,i}^{r_i} \varphi_{\bar{\lambda}}(x)|^p dx \right)^{1/p} \\ &\lesssim \left( \sum_{\bar{\lambda} \in \Delta_J} |a_{\bar{\lambda}}|^p \right)^{1/p} \omega_{r_i,i}(\varphi_{\bar{\lambda},J}, 2^{-k}, \Omega)_p. \end{aligned}$$

Thus we get by (5), (27), and Lemma 3.3 that

$$\begin{aligned} |f|_{B_{p,q}^{\bar{s}}} &\lesssim \sum_{i=1}^2 \|(2^{\ell_i s_i} \omega_{r_i,i}(f, 2^{-\ell_i}, \Omega)_p)_{\ell_i \geq 0}\|_{\ell_q} \\ &\lesssim \left( \sum_{\bar{\lambda} \in \Delta_J} |a_{\bar{\lambda}}|^p \right)^{1/p} \sum_{i=1}^2 \|(2^{\ell_i s_i} \omega_{r_i,i}(\varphi_{\bar{\lambda},J}, 2^{-\ell_i}, \Omega)_p)_{\ell_i \geq 0}\|_{\ell_q} \\ &\lesssim 2^{(j_1+j_2)(1/2-1/p)} \sum_{i=1}^2 2^{j_i s_i} \left( \sum_{\bar{\lambda} \in \Delta_J} |a_{\bar{\lambda}}|^p \right)^{1/p} \\ &\lesssim \sum_{i=1}^2 2^{j_i s_i} \|f\|_{p,\Omega}. \quad \blacksquare \end{aligned}$$

#### 4. CHARACTERIZATIONS BY APPROXIMATION SPACES

In the preceding section we proved Jackson and Bernstein estimates for two-parameter sequences of subspaces  $S_J$ . Now we consider specific subsequences of  $(S_J)_{J \in \mathbb{N}_0^2}$ , which depend on a particular anisotropy; i.e., we introduce anisotropy-dependent sequences of finite-dimensional subspaces. Let  $0 < a_2 < a_1 < 2$ ,  $a_1 + a_2 = 2$ ,  $s > 0$ ,  $s_1 := s/a_1$ ,  $s_2 := s/a_2$ . For  $n \in \mathbb{N}_0$  we set  $j_1(n) := n$ ,  $j_2(n) := \lfloor ns_1/s_2 \rfloor = \lfloor na_1/a_1 \rfloor$ ,

$$V_n := S_{(j_1(n), j_2(n))},$$



and  $V := V(a_1, a_2) := (V_n)_{n \in \mathbb{N}_0}$ . Clearly one has  $V_n \subset V_{n+1}$  for  $n \in \mathbb{N}_0$ . Furthermore, since  $2^{-s_2 j_2(n)} \lesssim 2^{-s_1 n}$ , we have for  $\bar{s} < M$ ,

$$\inf_{\chi \in V_n} \|f - \chi\|_p \lesssim 2^{-s_1 n} |f|_{B_{p,q}^{\bar{s}}}, \quad f \in B_{p,q}^{\bar{s}}, \quad (29)$$

if (26) holds, and for  $B_{p,q}^{\bar{s}}$ -smooth scaling functions it follows from (28) that

$$|f|_{B_{p,q}^{\bar{s}}} \lesssim 2^{s_1 n} \|f\|_p, \quad f \in V_n, \quad (30)$$

with constants independent of  $n \in \mathbb{N}_0$ , since  $2^{s_2 j_2(n)} \leq 2^{s_1 n}$ . Moreover, we set  $A_n := A_{(j_1(n), j_2(n))}$ .

With respect to  $V = (V_n)_{n \in \mathbb{N}}$  we introduce for  $a > 0$ ,  $0 < p < \infty$ , and  $0 < \mu < \infty$  the approximation spaces

$$A_{p,\mu}^\alpha := \left\{ f \in L_p(\Omega) \left| \left( \sum_{n=0}^{\infty} [2^{n\alpha} E_n(f)_p]^\mu \right)^{1/\mu} < \infty \right. \right\}, \quad (31)$$

where  $E_n(f)_p := \inf_{\chi \in V_n} \|f - \chi\|_p$ .

In our context the investigation of these approximation spaces, which, of course, turn out to be anisotropic Besov spaces, serves in some sense as a preliminary step toward a wavelet characterization. In fact, for  $1 \leq p < \infty$  the result of the following Theorem 4.1 immediately implies such a characterization; see Section 6. Only the case  $0 < p < 1$  enforces some further work; see Theorem 6.1.

On the other hand, the subsequent identity (35) is interesting by itself and provides various nontrivial consequences like interpolation and embedding results. Those results generalize known isotropic results (see [12]) and are, similar to the isotropic case, essentially a consequence of interpolation properties of the approximation spaces. For example, applying directly results from [11, 13] and the Jackson and Bernstein estimates, i.e., (29) and (30), we obtain

$$A_{p,\mu}^\alpha = (L_p, B_{p,q}^{\bar{s}})_{\alpha/s_1, \mu}, \quad 0 < \alpha < s_1. \quad (32)$$

For  $0 < p < 1$  what the right-hand side is does not seem to have been published. Therefore, the following identity (35) leads to a generalization of known interpolation results. Moreover, since for  $1 \leq p < \infty$  the interpolation spaces are already known (see [2, 28, 29]), Eq. (32) gives directly (35); i.e., nothing new has to be proved. Thus, Theorem 4.1 and its proof are only given because of the case  $0 < p < 1$ .

To begin with we remind the reader of discrete Hardy inequalities, which will be applied repeatedly in the following: First we introduce for  $a, q > 0$  sequence spaces  $\ell_q^\alpha$  equipped with the quasi-norms

$$\|(a_n)_{n \in \mathbb{N}_0}\|_{\ell_q^\alpha} := \begin{cases} \left( \sum_{n=0}^{\infty} [2^{n\alpha} |a_n|]^q \right)^{1/q}, & 0 < q < \infty, \\ \sup_n 2^{n\alpha} |a_n|, & q = \infty. \end{cases} \quad (33)$$

The discrete Hardy inequalities then read as follows (see [11]): For sequences  $(a_n)_{n \in \mathbb{N}_0}$  and  $(b_n)_{n \in \mathbb{N}_0}$  it holds that

$$\|(b_n)_{n \in \mathbb{N}_0}\|_{\ell_q^\alpha} \lesssim \|(a_n)_{n \in \mathbb{N}_0}\|_{\ell_q^\alpha}$$

if either

$$|b_n| \lesssim 2^{-n\lambda} \left( \sum_{j=0}^n [2^{j\lambda} |a_j|]^\mu \right)^{1/\mu}, \quad n \in \mathbb{N}_0, \quad (34)$$

or

$$|b_n| \lesssim \left( \sum_{j=n}^\infty |a_j|^\mu \right)^{1/\mu}, \quad n \in \mathbb{N}_0,$$

with  $\mu \leq q$  and in (34)  $\lambda > \alpha$ .

Additionally, we remind the reader of the following inequality: For  $\mu := \min(p, 1)$ ,  $0 < p \leq \infty$ , it holds that

$$\|f + g\|_p^\mu \leq \|f\|_p^\mu + \|g\|_p^\mu, \quad f, g \in L_p(\Omega).$$

**THEOREM 4.1.** *Let  $0 < p < \infty$  and  $0 < q \leq \infty$  and assume that the scaling functions are  $B_{p,q}^{\bar{s}}$ -smooth with  $\bar{s} \leq M$ . Then one has for  $\alpha > 0$  with  $\alpha(1, a_1/a_2) < \bar{s}$  that  $f \in B_{p,q}^{\alpha(1, a_1/a_2)}$  if and only if  $f \in A_{p,q}^\alpha$ ; i.e.,*

$$A_{p,q}^\alpha = B_{p,q}^{\alpha(1, a_1/a_2)}. \quad (35)$$

Moreover, one has with  $A_{-1}f := 0$  that

$$\begin{aligned} \|f\|_{B_{p,q}^{\alpha(1, a_1/a_2)}} &\sim \|f\|_p + \left( \sum_{n=0}^\infty [2^{n\alpha} E_n(f)_p]^q \right)^{1/q} \\ &\sim \left( \sum_{n=0}^\infty [2^{n\alpha} \|A_n(f) - A_{n-1}(f)\|_p]^q \right)^{1/q}. \end{aligned}$$

*Proof.* First we prove the second equivalence. The direction “ $\gtrsim$ ” follows from

$$E_n(f)_p \lesssim \|f - A_n(f)\|_p \lesssim \left[ \sum_{\ell=n}^\infty \|A_{n+1}(f) - A_n(f)\|_p^\mu \right]^{1/\mu},$$

where  $\mu := \min(1, p)$ , and a discrete Hardy inequality, since  $f = A_n(f) + \sum_{\ell=n}^\infty (A_{n+1}(f) - A_n(f))$ .

The direction “ $\lesssim$ ” follows directly from

$$\begin{aligned} \|A_{n+1}(f) - A_n(f)\|_p &\lesssim \|A_{n+1}(f) - f\|_p + \|f - A_n(f)\|_p \\ &\lesssim E_{n+1}(f)_p + E_n(f)_p. \end{aligned}$$

Thus it remains to investigate the first equivalence. To begin with we observe that (25) and (21) together with  $j_1(n) = n$  and  $j_2(n) = \lfloor na_2/a_1 \rfloor$  give

$$\begin{aligned} \left( \sum_{n=0}^{\infty} [2^{n\alpha} E_n(f)_p]^q \right)^{1/q} &\lesssim \left( \sum_{n=0}^{\infty} \left[ 2^{n\alpha} \sum_{i=1}^2 \omega_{m_i,i}(f, 2^{-j_i(n)})_p \right]^q \right)^{1/q} \\ &\lesssim \sum_{i=1}^2 \left( \sum_{n=0}^{\infty} [2^{n\alpha} \omega_{m_i,i}(f, 2^{-j_1(n)})_p]^q \right)^{1/q} \\ &\lesssim |f|_{B_{p,q}^{\alpha(1,a_1/a_2)}} \end{aligned}$$

since

$$\begin{aligned} \left( \sum_{n=0}^{\infty} [2^{n\alpha} \omega_{m_2,2}(f, 2^{-j_2(n)})_p]^q \right)^{1/q} &= \left( \sum_{\ell=0}^{\infty} \sum_{(a_1/a_2)\ell \leq n < (a_1/a_2)(\ell+1)} [2^{n\alpha} \omega_{m_2,2}(f, 2^{-\ell})_p]^q \right)^{1/q} \\ &\lesssim \left( \sum_{\ell=0}^{\infty} [2^{\ell(\alpha a_1/a_2)} \omega_{m_2,2}(f, 2^{-\ell})_p]^q \right)^{1/q} \\ &\lesssim |f|_{B_{p,q}^{\alpha(1,a_1/a_2)}}. \end{aligned}$$

That is, we have proved  $B_{p,q}^{\alpha(1,a_1/a_2)} \hookrightarrow A_{p,q}^{\alpha}$ .

Next, we show  $A_{p,q}^{\alpha} \hookrightarrow B_{p,q}^{\alpha(1,a_1/a_2)}$  by proving

$$\|(2^{s_i n} \omega_{m_i,i}(f, 2^{-j_i(n)})_p)_{n \in \mathbb{N}}\|_{\ell_q} \lesssim \|(2^{n\alpha} E_n(f)_p)_{n \in \mathbb{N}_0}\|_{\ell_q} + \|f\|_p, \quad (36)$$

for  $i = 1, 2$ . To this end we observe first that

$$\|(2^{s_i \ell} \omega_{m_i,i}(f, 2^{-j_i(\ell)})_p)_{\ell \in \mathbb{N}_0}\|_{\ell_q} \lesssim |f|_{B_{p,q}^{\bar{s}_i}} \lesssim 2^{s_1 n} \|f\|_p, \quad f \in V_n, \quad i = 1, 2,$$

implies for  $t > 0$  that

$$\omega_{m_1,1}(f, t)_p \lesssim \min(1, t2^n)^{s_1} \|f\|_p, \quad f \in V_n,$$

as well as

$$\omega_{m_2,2}(f, t)_p \lesssim \min(1, t2^{j_2(n)})^{s_2} \|f\|_p, \quad f \in V_n.$$

Thus we get for  $n \in \mathbb{N}_0$ , by (23), that

$$\begin{aligned} \omega_{m_1,1}(f, 2^{-n})_p^p &\leq \omega_{m_1,1}(A_0 f, 2^{-n})_p^p + \sum_{\ell=0}^{n-1} \omega_{m_1,1}(A_{\ell+1} f - A_{\ell} f, 2^{-n})_p^p \\ &\quad + \omega_{m_1,1}(f - A_n f, 2^{-n})_p^p \\ &\lesssim 2^{-ns_1 p} \|A_0 f\|_p^p + 2^{-ns_1 p} \sum_{\ell=0}^{n-1} 2^{s_1 \ell p} \|A_{\ell+1} f - A_{\ell} f\|_p^p + \|f - A_n f\|_p^p \\ &\lesssim 2^{-ns_1 p} \|f\|_p^p + 2^{-ns_1 p} \sum_{\ell=0}^n 2^{s_1 \ell p} E_{\ell}(f)_p^p, \end{aligned}$$

which gives by a Hardy inequality (36) for  $i = 1$ . Finally, we treat the case  $i = 2$ ; i.e., we have to show that for  $\beta := \alpha(a_1/a_2) < s_2$  it holds that

$$\sum_{n=0}^{\infty} 2^{\beta n q} \omega_{m_2, 2}(f, 2^{-n})_p^q \lesssim \|f\|_p^q + \sum_{n=0}^{\infty} 2^{n \alpha q} E_n(f)_p^q.$$

Since for  $\mu < \min(p, q)$

$$\left( \sum_{\ell=0}^n 2^{j_2(\ell) s_2 p} E_{\ell}(f)_p^p \right)^{1/p} \leq \left( \sum_{\ell=0}^n 2^{j_2(\ell) s_2 \mu} E_{\ell}(f)_p^{\mu} \right)^{1/\mu},$$

we obtain analogously to the above by (23) that

$$\omega_{m_2, 2}(f, 2^{-n})_p^p \lesssim 2^{-n s_2 p} \|f\|_p^p + 2^{-n s_2 p} \left( \sum_{\ell=0}^n 2^{j_2(\ell) s_2 \mu} E_{\ell}(f)_p^{\mu} \right)^{p/\mu}.$$

Therefore, by Hölder inequalities and by transferring ideas from the isotropic setting again (see, e.g., [27]) we see that

$$\begin{aligned} & \sum_{n=0}^{\infty} 2^{\beta n q} \omega_{m_2, 2}(f, 2^{-n})_p^q \\ & \lesssim \sum_{n=0}^{\infty} 2^{\beta n q} \left( 2^{-n s_2 q} \|f\|_p^q + 2^{-n s_2 q} \left( \sum_{\ell=0}^n 2^{j_2(\ell) s_2 \mu} E_{\ell}(f)_p^{\mu} \right)^{q/\mu} \right) \\ & \lesssim \sum_{n=0}^{\infty} 2^{n q(\beta - s_2)} \|f\|_p^q + \sum_{n=0}^{\infty} 2^{n q(\beta - s_2)} \left( \sum_{\ell=0}^n 2^{j_2(\ell) s_2 \mu} E_{\ell}(f)_p^{\mu} \right)^{q/\mu} \\ & \lesssim \|f\|_p^q + \sum_{n=0}^{\infty} 2^{n q(\beta - s_2)} \left( \sum_{\ell=0}^n 2^{j_2(\ell)(s_2 - \beta)(\mu - \mu^2/q)} 2^{j_2(\ell)(s_2 - \beta)\mu^2/q} 2^{j_2(\ell)\beta\mu} E_{\ell}(f)_p^{\mu} \right)^{q/\mu} \\ & \lesssim \|f\|_p^q + \sum_{n=0}^{\infty} 2^{n q(\beta - s_2)} \left( \sum_{\ell=0}^n 2^{j_2(\ell)(s_2 - \beta)\mu} \right)^{q/\mu - 1} \sum_{\ell=0}^n 2^{j_2(\ell)(s_2 - \beta)\mu} 2^{j_2(\ell)\beta q} E_{\ell}(f)_p^q \\ & \lesssim \|f\|_p^q + \sum_{n=0}^{\infty} 2^{n(\beta - s_2)\mu} \sum_{\ell=0}^n 2^{j_2(\ell)(s_2 - \beta)\mu} 2^{j_2(\ell)\beta q} E_{\ell}(f)_p^q \\ & \lesssim \|f\|_p^q + \sum_{\ell=0}^{\infty} \left( \sum_{j \geq \ell} 2^{j(\beta - s_2)\mu} \right) 2^{j_2(\ell)(s_2 - \beta)\mu} 2^{j_2(\ell)\beta q} E_{\ell}(f)_p^q \\ & \lesssim \|f\|_p^q + \sum_{\ell=0}^{\infty} 2^{j_2(\ell)\beta q} E_{\ell}(f)_p^q \\ & \lesssim \|f\|_p^q + \sum_{n=0}^{\infty} 2^{\alpha n q} E_n(f)_p^q. \blacksquare \end{aligned}$$

We remark that an analog to (35) holds for  $p = \infty$ . But treating this case would have required some additional care; e.g., some estimates in Sections 3.2 and 3.3 are not true

for  $f \in L_\infty(\Omega)$  but are for  $f \in C(\overline{\Omega})$ . To streamline our presentation we neglected such “details.”

Further, we note that  $B_{p,q}^{\bar{s}}$ -smooth scaling functions with  $0 < q \leq \infty$  are for suitable  $\epsilon > 0$  also  $B_{p,q}^{\bar{s}-(\epsilon,\epsilon)}$ -smooth. Thus, instead of (35) we may write

$$A_{p,r}^\alpha = B_{p,r}^{\alpha(1,a_1/a_2)}, \quad 0 < r \leq \infty.$$

## 5. INTERPOLATION AND EMBEDDING RESULTS

In this section we present some consequences from the characterization result in Section 4. In particular, we investigate interpolation and embedding properties.

First we observe that for  $a_1 = a_2 = 1$ , i.e.,  $s_1 = s_2 = s > 0$ , and for  $0 < p < \infty$ ,  $0 < q \leq \infty$  holds

$$B_{p,q}^{(s,s)} = B_{p,q}^s, \quad (37)$$

since the approximation spaces are in this case already identified as isotropic Besov spaces; cf. [4, 5, 13, 25, 27]. That is, the isotropic Besov spaces are specific anisotropic Besov spaces.

Another immediate consequence concerns interpolation properties for anisotropic Besov spaces: Eqs. (32) and (35) imply for  $0 < r \leq \infty$  that

$$(L_p, B_{p,q}^{\bar{s}})_{\alpha/s_1, r} = B_{p,r}^{\alpha(1,a_1/a_2)}, \quad 0 < \alpha < s_1, \quad (38)$$

which generalizes results presented in [28, 29]. We remark that at least for  $1 \leq p < \infty$  (38) may be concluded from results in [28, 29] as follows: For  $\bar{\ell} \in \mathbb{N}_0^2$  anisotropic Sobolev spaces are defined by

$$W_p^{\bar{\ell}}(\Omega) := \left\{ f \in L_p(\Omega) \left| D^\alpha f \in L_p(\Omega), \sum_{j=1}^2 \frac{\alpha_j}{\ell_j} \leq 1 \right. \right\}.$$

In [29] it is proved that for  $0 < \theta < 1$  and  $\bar{s} = \theta \bar{\ell}$ ,

$$B_{p,q}^{\bar{s}}(\Omega) = (L_p(\Omega), W_p^{\bar{\ell}}(\Omega))_{\theta, q}$$

holds. Thus, a standard reiteration theorem provides (38) for  $1 \leq p < \infty$ .

Moreover, we note that in [28] one also finds the identity

$$B_{p,q}^{\bar{s}}(\Omega) = (L_p(\Omega), W_p^{(\ell_1, 0)}(\Omega))_{\theta, q} \cap (L_p(\Omega), W_p^{(0, \ell_2)}(\Omega))_{\theta, q}, \quad (39)$$

for  $\bar{s} = \theta(\ell_1, \ell_2)$ ,  $1 \leq p < \infty$ , and  $1 \leq q \leq \infty$ .

We remark that by arguments which are similar to those above one can prove

$$\omega_{\ell_1, 1}(f, t, \Omega)_p \sim K(f, t^{\ell_1}, L_p(\Omega), W_p^{(\ell_1, 0)}(\Omega)) \quad (40)$$

and analogously that

$$\omega_{\ell_2, 2}(f, t)_p \sim K(f, t^{\ell_2}, L_p(\Omega), W_p^{(0, \ell_2)}(\Omega)), \quad (41)$$

where  $K$  denotes the usual  $K$ -functional with respect to Banach spaces. The relations (40) and (41) give then an alternative proof for the fact (39).

Next we give an anisotropic generalization of the isotropic interpolation result presented in [12]. This generalization is in some sense sharp because of the anisotropic counterexample given in [29].

**THEOREM 5.1.** *For  $\alpha_0, \alpha_1 > 0$ ,  $0 < p_0, p_1, q_0, q_1 < \infty$ , and  $0 < \Theta < 1$  with  $1/q := \Theta/q_0 + (1 - \Theta)/q_1$ ,  $1/p := \Theta/p_0 + (1 - \Theta)/p_1$ ,*

$$(B_{p_0, q_0}^{\alpha_0 \bar{s}}, B_{p_1, q_1}^{\alpha_1 \bar{s}})_{\Theta, q} = B_{p, q}^{\alpha \bar{s}} \quad (42)$$

*holds with  $\alpha = \Theta\alpha_0 + (1 - \Theta)\alpha_1$ , provided  $p = q$ .*

*Proof.* The following sketch of a proof for (42) is essentially a modification of Chapter 6 in [12]: For a quasi-Banach space  $(X, \|\cdot\|_X)$  and  $a > 0$ ,  $0 < q \leq \infty$  we introduce the quasi-Banach space  $\ell_q^\alpha(X)$ , which consists, analogously to (33), of all sequences  $(a_n)_{n \in \mathbb{N}_0} \subset X$  such that

$$\|(a_n)_{n \in \mathbb{N}_0}\|_{\ell_q^\alpha(X)} := \|(2^{n\alpha} \|a_n\|_X)_{n \in \mathbb{N}_0}\|_{\ell_q} < \infty.$$

In [11] the following interpolation property is proved: If  $X_0, X_1$  are a pair of quasi-Banach spaces and if  $\alpha := (1 - \theta)\alpha_0 + \theta\alpha_1$ ,  $1/q = (1 - \theta)/q_0 + \theta/q_1$ , where  $0 < q_0, q_1 \leq \infty$ ,  $\alpha_0, \alpha_1 > 0$ , and  $0 < \theta < 1$ , one has

$$(\ell_{q_0}^{\alpha_0}(X_0), \ell_{q_1}^{\alpha_1}(X_1))_{\theta, q} = \ell_q^\alpha((X_0, X_1)_{\theta, q}). \quad (43)$$

In order to prove (42) we set  $X_0 := L_{p_0}$  and  $X_1 := L_{p_1}$ . Next we define a mapping, which relates  $B_{p_i, q_i}^{\alpha_i \bar{s}}$  to  $\ell_{q_i}^{\alpha_i}(X_i)$ . To this end we note that the approach in [12] applies the following observation from [14]: Let  $I$  and  $J$  denote cubes, i.e., quadratic rectangles, in  $\mathbb{R}^2$  with  $I \subset J$  and  $|J| \leq a|I|$ ; then there is a constant  $c > 0$  depending on  $a$  and independent of  $0 < p < \infty$  such that for any polynomial  $P$  of coordinate degree  $< r$  one has

$$\|P\|_{p, J} \leq c \|P\|_{p, I}. \quad (44)$$

This independence of  $p$  is no longer true if  $I$  and  $J$  are allowed to be arbitrary dyadic rectangles in  $\mathcal{D}$ . Thus the constant in (44) depends on  $p$ . Moreover, if with respect to a cube  $I$  and a function  $f \in L_\rho(I)$ ,  $\rho > 0$ , a near-best polynomial approximation of coordinate degree  $< r$  is denoted by  $Q_\rho$ , it is shown in [12] that this turns out to be also a near-best  $L_p(I)$  approximation on the cube  $I$  for all  $p \geq \rho$ . Here again, the anisotropy in the admitted dyadic rectangles destroys the underlying basic estimates for polynomials which provide this property; i.e., we are not able to prove independent of  $p \geq p$ ,  $\rho > 0$  fixed, that near-best approximations on  $\square \in \mathcal{D}$  are also near-near-best approximations on  $\tilde{\square}$ . In both problems we can resolve the situation by fixing  $p_0$  and  $p_1$  in the above situation.

**LEMMA 5.1.** *Let  $0 < p_0 \leq p_1 < \infty$ . Then there is a constant  $c = c(p_0, p_1, M)$  such that*

$$\|P\|_{p_1, \square} \leq c |\square|^{(1/p_1 - 1/p_0)} \|P\|_{p_0, \square}, \quad P \in \Pi^M, \square \in \mathcal{D}.$$

*Proof.* First we note that the equivalence of quasi-norms on  $\Pi^M$  gives

$$\|P\|_{p_1, \Omega} \leq c(p_0, p_1, M) \|P\|_{p_0, \Omega}.$$

For an arbitrary but fixed  $\square \in \mathcal{D}$  we denote the affine mapping which maps  $\square$  onto  $\Omega$  by  $C$  and set  $\tilde{P} := P \circ C^{-1}$ . Then we get

$$\begin{aligned} \|P\|_{p_1, \square}^{p_1} &= |\square| \|\tilde{P}\|_{p_1, \Omega}^{p_1} \leq c(p_0, p_1, M) |\square| \|\tilde{P}\|_{p_0, \Omega}^{p_1} \\ &= c(p_0, p_1, M) |\square|^{1-p_1/p_0} \|P\|_{p_0, \square}^{p_1}. \blacksquare \end{aligned}$$

LEMMA 5.2. Let  $0 < p_0 \leq p_1 < \infty$  and let  $Q_{p_0} \in \Pi^M$  be a near-best  $L_{p_0}(\square)$ -approximation,  $\square \in \mathcal{D}$  for  $f \in L_{p_1}(\square)$ ; i.e., for a constant  $A \geq 1$

$$\|f - Q_{p_0}\|_{p_0, \square} \leq A \inf_{Q \in \Pi^M} \|f - Q\|_{p_0, \square}$$

holds. Then there is a constant  $c = c(p_0, p_1, M)$  such that

$$\|f - Q_{p_0}\|_{p_1, \square} \leq cA^2 \inf_{Q \in \Pi^M} \|f - Q\|_{p_1, \square}.$$

*Proof.* Let  $Q_{p_1} \in \Pi^M$  denote a near-best  $L_{p_1}(\square)$ -approximation to  $f$ . Then by Lemma 5.1

$$\begin{aligned} \|f - Q_{p_0}\|_{p_1, \square} &\leq \|f - Q_{p_1}\|_{p_1, \square} + \|Q_{p_1} - Q_{p_0}\|_{p_1, \square} \\ &\leq A \inf_{Q \in \Pi^M} \|f - Q\|_{p_1, \square} + c(p_0, p_1, M) |\square|^{1/p_1-1/p_0} \|Q_{p_1} - Q_{p_0}\|_{p_0, \square} \end{aligned}$$

follows. The Hölder inequality gives

$$\begin{aligned} \|Q_{p_1} - Q_{p_0}\|_{p_0, \square} &\leq \|f - Q_{p_1}\|_{p_0, \square} + \|f - Q_{p_0}\|_{p_0, \square} \leq (1+A) \|f - Q_{p_1}\|_{p_0, \square} \\ &\leq (1+A) |\square|^{1/p_0-1/p_1} \|f - Q_{p_1}\|_{p_1, \square}, \end{aligned}$$

and consequently we obtain

$$\begin{aligned} \|f - Q_{p_0}\|_{p_1, \square} &\leq (A + A(A+1)c(p_0, p_1)) \inf_{Q \in \Pi^M} \|f - Q\|_{p_1, \square} \\ &\leq A^2(1 + 2c(p_0, p_1, M)) \inf_{Q \in \Pi^M} \|f - Q\|_{p_1, \square}. \blacksquare \end{aligned}$$

Lemma 5.2 allows us to introduce for fixed  $0 < p_0, p_1 < \infty$  and suitable functions  $f$  for each dyadic rectangle  $\square \in \mathcal{D}$  a polynomial in  $\Pi^M$ , which is a near-best approximation with respect to  $L_{p_0}(\square)$  and  $L_{p_1}(\square)$  with a common constant independent of  $\square$ . Furthermore, this common near-best approximation on  $\square$  is also a near-best approximation on the related dyadic rectangles  $\tilde{\square}$  with constants independent of  $\square$ , though, contrary to [12], dependent on  $p_0$  and  $p_1$ .

Proceeding with the proof of (42), let us assume for the following without any restriction that  $0 < p_0 \leq p_1 < \infty$ : If we start with local near-best approximations on  $\square$  for  $f \in L_{p_0}(\Omega)$  with respect to  $L_{p_0}(\square)$ , we obtain via  $A_n(f)$  and  $a_n(f) := A_n(f) - A_{n-1}(f)$ ,  $n \in \mathbb{N}_0$ , a mapping on  $L_{p_0}(\Omega)$  defined by  $Af := (a_n(f))_{n \in \mathbb{N}}$ , such that  $f \in B_{p_i, q_i}^{\alpha_i \tilde{s}}$  if and

only if  $Af \in \ell_{q_i}^{\alpha_i}(L_{p_i})$  and also  $\|f\|_{B_{p_i, q_i}^{\alpha_i \bar{s}}} \sim \|Af\|_{\ell_{q_i}^{\alpha_i}(L_{p_i})}$ ; cf. Theorem 4.1. Thus, if we could prove for  $t > 0$  that

$$K(f, t, B_{p_0, q_0}^{\alpha_0 \bar{s}}, B_{p_1, q_1}^{\alpha_1 \bar{s}}) \sim K(Af, t, \ell_{q_0}^{\alpha_0}(L_{p_0}), \ell_{q_1}^{\alpha_1}(L_{p_1})), \quad (45)$$

then we would get by standard interpolation arguments that  $f \in (B_{p_0, q_0}^{\alpha_0 \bar{s}}, B_{p_1, q_1}^{\alpha_1 \bar{s}})_{\theta, q}$  if and only if  $Af \in (\ell_{q_0}^{\alpha_0}(L_{p_0}), \ell_{q_1}^{\alpha_1}(L_{p_1}))_{\theta, q}$  for  $0 < \theta < 1$ ,  $0 < q \leq \infty$ , together with equivalent quasi-norms. Thus, applying (43) and finally

$$(L_{p_0}(\Omega), L_{p_1}(\Omega))_{\theta, q} = L_{p, q}(\Omega), \quad \frac{1}{p} = \frac{\theta}{p_0} + \frac{1-\theta}{p_1},$$

where  $L_{p, q}(\Omega)$  denotes Lorentz spaces, we end up with (42). Thus it remains to prove (45). But taking into account the above modifications and our results in Sections 3.1 and 3.2, the arguments from [12, Chap. 6] can now be transferred verbatim. ■

Next we prove an embedding result for the anisotropic spaces under consideration. To this end we note that Lemma 5.1 implies for  $p \leq r$ ,  $P \in \Pi^M$ , and  $\square \in \mathcal{D}_J$  that

$$\|P\|_{r, \square} \leq c(p, r) 2^{(j_1+j_2)(1/r-1/p)} \|P\|_{p, \square}.$$

Thus we get for  $S \in \Pi_J^M$ ,

$$\begin{aligned} \|S\|_{r, \Omega}^r &= \sum_{\square \in \mathcal{D}_J} \|S\|_{r, \square}^r \lesssim 2^{(j_1+j_2)(1-p/r)} \sum_{\square \in \mathcal{D}_J} \|S\|_{p, \square}^r \\ &\lesssim 2^{(j_1+j_2)(1-p/r)} \left( \sum_{\square \in \mathcal{D}_J} \|S\|_{p, \square}^p \right)^{r/p}, \end{aligned}$$

since  $p/r \leq 1$ , which gives

$$\|S\|_{r, \Omega} \lesssim 2^{(j_1+j_2)(1/p-1/r)} \|S\|_{p, \Omega}. \quad (46)$$

In particular, for  $j_1(n) = n$ , and  $j_2(n) = \lfloor na_2/a_1 \rfloor$ ,  $n \in \mathbb{N}_0$ , and  $1/p - 1/r = s/2 = s_1 s_2 / (s_1 + s_2) > 0$  (46) yields for  $v \in V_n$

$$\|v\|_{r, \Omega} \lesssim 2^{(n+ns_1/s_2)(1/p-1/r)} \|v\|_{p, \Omega} = 2^{ns_1} \|v\|_{p, \Omega}.$$

**THEOREM 5.2.** *For  $\bar{s} = (s_1, s_2) > 0$  and  $0 < p < r < \infty$  with  $1/p - 1/r = s/2$ ,  $B_{p, r}^{\bar{s}} \hookrightarrow L_r(\Omega)$  holds; i.e.,*

$$\|f\|_r \lesssim \|f\|_{B_{p, r}^{\bar{s}}}, \quad f \in B_{p, r}^{\bar{s}}. \quad (47)$$

*Proof.* Let us first prove the weaker embedding

$$B_{p, \mu}^{\bar{s}} \hookrightarrow L_r(\Omega) \quad (48)$$

with  $\mu := \min(1, r)$ : We fix  $\bar{s} > 0$  and  $0 < p < r < \infty$  and choose a suitable multiresolution such that Theorem 4.1 is applicable with respect to the anisotropic



Besov space  $B_{p,r}^{\bar{s}}$ . Then for  $f \in B_{p,r}^{\bar{s}}$  we obtain in  $L_p(\Omega)$ ,  $f = \sum_{n=0}^{\infty} A_n f - A_{n-1} f$  ( $A_{-1} f := 0$ ), and consequently for  $\mu = \min(1, r)$ ,

$$\|f\|_r \leq \left( \sum_{n=0}^{\infty} \|A_n f - A_{n-1} f\|_r^{\mu} \right)^{1/\mu} \lesssim \left( \sum_{n=0}^{\infty} (2^{ns_1} \|A_n f - A_{n-1} f\|_p)^{\mu} \right)^{1/\mu} \lesssim \|f\|_{B_{p,\mu}^{\bar{s}}}.$$

If  $r > 1$ , the embedding (48) states only  $B_{p,1}^{\bar{s}} \hookrightarrow L_r(\Omega)$  instead of  $B_{p,r}^{\bar{s}} \hookrightarrow L_r(\Omega)$ . We remark that for our proof of the wavelet characterization in Theorem 6.1 this would be sufficient.

The more general embedding (47) can be shown by applying the interpolation result (38) and generalizing a further idea from [12] as follows: For  $r > 1$  we choose  $r_i$ ,  $i = 0, 1$ , such that  $1 \leq r_0 < r < r_1 < \infty$ , and set  $\alpha_1 = (1/p - 1/r_i)(2/a_1)$ . Then we get by (48) that

$$\|f\|_{r_i} \lesssim \|f\|_{B_{p,1}^{\alpha_1(1,a_1/a_2)}},$$

and for  $\theta \in (0, 1)$  with  $1/r = \theta/r_0 + (1 - \theta)/r_1$  it follows by (38) and interpolation theory that

$$\|f\|_r \lesssim \|f\|_{B_{p,r}^{\alpha(1,a_1/a_2)}},$$

where we used  $\alpha = (1/p - 1/r)(2/a_1) = 2/pa_1 + (\theta/r_0 + (1 - \theta)/r_1)(2/a_1) = \theta\alpha_0 + (1 - \theta)\alpha_1$ . ■

We note that for  $p \geq 1$  Theorem 5.2 is known; see, e.g., the survey given in [23].

## 6. WAVELET CHARACTERIZATIONS

Now we show wavelet characterizations for the anisotropic Besov spaces under consideration. As above we fix the anisotropy by choosing  $0 < a_2 \leq a_1 < 2$  with  $a_1 + a_2 = 2$  and consider again a related sequence of finite-dimensional subspaces  $V = V(a_1, a_2) = (V_n)_{n \in \mathbb{N}}$ .

To formulate the wavelet characterization results we introduce corresponding anisotropy-dependent index sets  $\Lambda_n$ ,  $n \in \mathbb{N}_0$ , which are related to the complements of  $V_n$  in  $V_{n+1}$ : If for  $n \in \mathbb{N}_0$ ,  $j_2(n+1) = j_2(n)$  holds, we set

$$\Lambda_{n+1} := \{(\lambda_1, \lambda_2) \mid \lambda_1 \in \nabla_{n+1}^{(1)}, \lambda_2 \in \Delta_{j_2(n)}^{(2)}\}.$$

If  $j_2(n+1) = j_2(n) + 1$ , then

$$\begin{aligned} \Lambda_{n+1} &:= \{(\lambda_1, \lambda_2) \mid \lambda_1 \in \nabla_{n+1}^{(1)}, \lambda_2 \in \Delta_{j_2(n)}^{(2)}, \\ &\quad \lambda_1 \in \Delta_n^{(1)}, \lambda_2 \in \nabla_{j_2(n)+1}^{(2)}, \\ &\quad \lambda_1 \in \nabla_{n+1}^{(1)}, \lambda_2 \in \nabla_{j_2(n)+1}^{(2)}\}. \end{aligned}$$

Moreover, we set  $\Lambda_0 := \nabla_0^{(1)} \times \nabla_0^{(2)}$ .

For  $\bar{\lambda} \in \Lambda_n$  we denote the corresponding biorthogonal tensor-product wavelet basis functions by  $\psi_{\bar{\lambda}}$  and  $\tilde{\psi}_{\bar{\lambda}}$ , respectively; i.e., e.g., if  $\bar{\lambda} = (\lambda_1, \lambda_2) \in \Delta_n^{(1)} \times \nabla_{j_2(n)+1}^{(2)}$  then  $\psi_{\bar{\lambda}} = \varphi_{\lambda_1}^{(1)} \otimes \psi_{\lambda_2}^{(2)}$ , or if  $\bar{\lambda} = (\lambda_1, \lambda_2) \in \nabla_{n+1}^{(1)} \times \nabla_{j_2(n)+1}^{(2)}$  then  $\psi_{\bar{\lambda}} = \psi_{\lambda_1}^{(1)} \otimes \psi_{\lambda_2}^{(2)}$ .

To begin with we discuss wavelet characterizations for  $1 \leq p < \infty$ . It turns out that they can, more or less, be deduced directly from the results in Section 4: Let  $1 \leq p < \infty$ ,  $0 < q \leq \infty$ , and assume that the scaling functions are  $B_{p,q}^{\bar{s}}$ -smooth with  $\bar{s} \leq M$ . Then, replacing  $A_n$  with  $P_n$  and setting  $P_{-1} := 0$ , the proof of Theorem 4.1 implies for  $\alpha > 0$  and  $\alpha(1, a_1/a_2) < \bar{s}$  that

$$\|f\|_{B_{p,q}^{\alpha(1,a_1/a_2)}} \sim \left( \sum_{n=0}^{\infty} [2^{n\alpha} \|(P_n - P_{n-1})f\|_p]^q \right)^{1/q}, \quad f \in B_{p,q}^{\alpha(1,a_1/a_2)}, \quad (49)$$

and  $f \in L_p(\Omega)$  lies in  $B_{p,q}^{\alpha(1,a_1/a_2)}$  if and only if the right-hand side in (49) is finite. On the other hand, the above definitions imply  $(P_n - P_{n-1})f = \sum_{\tilde{\lambda} \in \Lambda_n} \langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle \psi_{\tilde{\lambda}}$  for  $n \in \mathbb{N}$ , and our assumptions concerning FWT provide taking into account

$$\left\| \sum_{\tilde{\lambda} \in \Delta_j} a_{\tilde{\lambda}} \varphi_{\tilde{\lambda}} \right\|_{p,\Omega} \sim 2^{(j_1+j_2)(1/2-1/p)} \|(a_{\tilde{\lambda}})_{\tilde{\lambda} \in \Delta_j}\|_{\ell_p},$$

that

$$\|(P_n - P_{n-1})f\|_p \sim 2^{j(n)(1/2-1/p)} \|(\langle f, \tilde{\psi}_{\tilde{s}} \rangle)_{\tilde{\lambda} \in \Lambda_n}\|_{\ell_p},$$

where  $j(n) := j_1(n) + j_2(n)$ . Thus, for  $1 \leq p < \infty$ , we end up with

$$f \in B_{p,q}^{\alpha(1,a_1/a_2)} \quad \text{iff} \quad \left( \sum_{n=0}^{\infty} \left[ 2^{n\alpha+j(n)(1/2-1/p)} \left( \sum_{\tilde{\lambda} \in \Lambda_n} |\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle|^p \right)^{1/p} \right]^q \right)^{1/q} < \infty$$

with equivalent (quasi-)norms.

Already the study of isotropic Besov spaces has shown that the analog to the last statement is in general not true for all  $0 < p < 1$ ; see [24]. An obvious problem is that, generally,  $\langle f, \tilde{\varphi}_{\tilde{\lambda}} \rangle$  as well as  $\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle$  is not defined without assuming some additional smoothness. Furthermore, replacing in the above arguments  $P_n$  with  $A_n$ , one has in general

$$A_n f - A_{n-1} f \neq \sum_{\tilde{\lambda} \in \Lambda_n} \langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle \psi_{\tilde{\lambda}}.$$

Thus, even under suitable regularity assumptions Theorem 4.1 would not imply directly an appropriate wavelet characterization. But we shall see that an application of Theorem 5.2 will lead to the desired result.

**THEOREM 6.1.** *Let  $0 < p < \infty$  and  $0 < q \leq \infty$  and assume that the scaling functions are  $B_{p,q}^{\bar{s}}$ -smooth with  $\bar{s} \leq M$ . Then for  $\alpha > 0$  with  $\alpha(1, a_1/a_2) < \bar{s}$  and  $\frac{1}{\min(1,p)} - 1 < \alpha a_1/2$  one has  $f \in B_{p,q}^{\alpha(1,a_1/a_2)}$  if and only if  $f = \sum_{n \in \mathbb{N}_0} \sum_{\tilde{\lambda} \in \Lambda_n} f_{\tilde{\lambda}} \psi_{\tilde{\lambda}} \in L_p(\Omega)$  with  $f_{\tilde{\lambda}} := \langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle$  and*

$$\left( \sum_{n \in \mathbb{N}_0} \left[ 2^{n\alpha+j(n)(1/2-1/p)} \left( \sum_{\tilde{\lambda} \in \Lambda_n} |f_{\tilde{\lambda}}|^p \right)^{1/p} \right]^q \right)^{1/q} < \infty. \quad (50)$$

Furthermore, (50) defines an equivalent (quasi-)norm.

*Proof.* Because of our above remarks it remains to consider  $0 < p < 1$ . Since for  $Q_n f := (P_n - P_{n-1})f$ ,  $n \in \mathbb{N}_0$ ,

$$\|Q_n f\|_p \sim 2^{j(n)(1/2-1/p)} \left( \sum_{\tilde{\lambda} \in \Delta_{n,j_2(n)}} |\langle Q_n f, \tilde{\varphi}_{\tilde{\lambda}} \rangle|^p \right)^{1/p}$$

holds (see Lemma 3.3), one has

$$\|Q_n f\|_p \sim 2^{j(n)(1/2-1/p)} \left( \sum_{\tilde{\lambda} \in \Lambda_n} |\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle|^p \right)^{1/p}.$$

Thus it is sufficient to prove

$$f \in B_{p,q}^{\alpha(1,a_1/a_2)} \quad \text{iff } f \in L_p(\Omega) \text{ and } \left( \sum_{n=0}^{\infty} (2^{n\alpha} \|Q_n f\|_p)^q \right)^{1/q} < \infty.$$

To this end we introduce for  $0 < q \leq \infty$ ,

$$S_{p,q}^{\alpha} := \left\{ f \in L_r(\Omega) \left| \left( \sum_{n=0}^{\infty} (2^{n\alpha} \|Q_n f\|_p)^q \right)^{1/q} < \infty \right. \right\}, \quad (51)$$

where  $1/r := 1/p - \alpha a_1/2 < 1$ .

In the first step we show that  $B_{p,p}^{\alpha(1,a_1/a_2)} \hookrightarrow S_{p,\infty}^{\alpha}$ : Since by definition it follows that  $B_{p,p}^{\alpha(1,a_1/a_2)} \hookrightarrow B_{p,1}^{\alpha(1,a_1/a_2)}$  one has by Theorem 5.2 that  $B_{p,p}^{\alpha(1,a_1/a_2)} \hookrightarrow L_r(\Omega)$ . Thus the terms  $\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle$  are defined and Hölder's inequality and biorthogonality give for  $n \in \mathbb{N}$ ,

$$|\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle| \lesssim 2^{-j(n)(1/2-1/r)} \inf_{P \in \Pi^M} \|f - P\|_{r, \text{supp}(\tilde{\psi}_{\tilde{\lambda}})}, \quad \tilde{\lambda} \in \Lambda_n.$$

With respect to  $\Omega = [0, 1]^2$  we proved for  $1/p - 1/r = s/2$  that

$$\|f\|_{r,\Omega} \lesssim \|f\|_{B_{p,1}^{\alpha(1,a_1/a_2)}}, \quad f \in B_{p,1}^{\alpha(1,a_1/a_2)},$$

which implies by Theorem 3.1 that

$$\inf_{P \in \Pi^M} \|f - P\|_{r,\Omega} \lesssim \inf_{P \in \Pi^M} \|f - P\|_{p,\Omega} + |f|_{B_{p,1}^{\alpha(1,a_1/a_2)}} \lesssim |f|_{B_{p,1}^{\alpha(1,a_1/a_2)}}. \quad (52)$$

Defining  $\tilde{\Omega}_{\tilde{\lambda}} := \text{supp } \tilde{\psi}_{\tilde{\lambda}}$ ,  $\tilde{\lambda} \in \Lambda_n$ , we get

$$\|f\|_{r,\tilde{\Omega}_{\tilde{\lambda}}} \sim 2^{-j(n)/r} \|\tilde{f}\|_{r,\Omega},$$

where  $\tilde{f}$  denotes the translated and rescaled version of  $f|_{\tilde{\Omega}_{\tilde{\lambda}}}$  on  $\Omega$ .

Next we note that for  $\ell \in \mathbb{N}_0$ ,

$$\begin{aligned} \omega_{m_1,1}(\tilde{f}, 2^{-\ell}, \Omega)_p &\lesssim 2^{j(n)/p} \omega_{m_1,1}(f, 2^{-(\ell+n)}, \tilde{\Omega}_{\tilde{\lambda}})_p \quad \text{and} \\ \omega_{m_2,2}(\tilde{f}, 2^{-j_2(\ell)}, \Omega)_p &\lesssim 2^{j(n)/p} \omega_{m_2,2}(f, 2^{-j_2(\ell+n)}, \tilde{\Omega}_{\tilde{\lambda}})_p \end{aligned}$$

hold, which implies by (52) and by taking into account that  $2^{j(n)(1/p-1/r)} \sim 2^{\alpha n}$  that

$$\begin{aligned}
 \inf_{P \in \Pi^M} \|f - P\|_{r, \tilde{\Omega}_{\tilde{\lambda}}} &\lesssim 2^{-j(n)/p} \inf_{P \in \Pi^M} \|\tilde{f} - P\|_{r, \Omega} \\
 &\lesssim 2^{-j(n)/r} \sum_{\ell=0}^{\infty} 2^{\ell\alpha} \sum_{i=1}^2 \omega_{m_i, i}(\tilde{f}, 2^{-j_i(\ell)}, \Omega)_p \\
 &\lesssim \sum_{\ell=0}^{\infty} 2^{\alpha(n+\ell)} \sum_{i=1}^2 \omega_{m_i, i}(f, 2^{j_i(n+\ell)}, \tilde{\Omega}_{\tilde{\lambda}})_p \\
 &\lesssim |f|_{B_{p,1}^{\alpha(1,a_1/a_2)}(\tilde{\Omega}_{\tilde{\lambda}})}.
 \end{aligned}$$

Next, by the summing property of the averaged moduli follows

$$\begin{aligned}
 &\sum_{\tilde{\lambda} \in \Lambda_n} |f|_{B_{p,p}^{\alpha(1,a_1/a_2)}(\tilde{\Omega}_{\tilde{\lambda}})}^p \\
 &\lesssim \sum_{\lambda \in \Lambda_n} \left( \sum_{j=0}^{\infty} [2^{\alpha j} \tilde{\omega}_{m_1,1}(f, 2^{-j}, \tilde{\Omega}_{\tilde{\lambda}})_p]^p + \sum_{j=0}^{\infty} [2^{\alpha a_1/a_2 j} \tilde{\omega}_{m_2,2}(f, 2^{-j}, \tilde{\Omega}_{\tilde{\lambda}})_p]^p \right) \\
 &\lesssim \sum_{j=0}^{\infty} [2^{\alpha j} \tilde{\omega}_{m_1,1}(f, 2^{-j}, \Omega)_p]^p + \sum_{j=0}^{\infty} [2^{\alpha a_1/a_2 j} \tilde{\omega}_{m_2,2}(f, 2^{-j}, \Omega)_p]^p \\
 &\lesssim \sum_{j=0}^{\infty} [2^{\alpha j} \omega_{m_1,1}(f, 2^{-j}, \Omega)_p]^p + \sum_{j=0}^{\infty} [2^{\alpha a_1/a_2 j} \omega_{m_2,2}(f, 2^{-j}, \Omega)_p]^p \\
 &\lesssim |f|_{B_{p,p}^{\alpha(1,a_1/a_2)}(\Omega)}^p,
 \end{aligned}$$

which provides

$$\begin{aligned}
 2^{j(n)(1/2-1/p)} \left( \sum_{\tilde{\lambda} \in \Lambda_n} |\langle f, \tilde{\psi}_{\tilde{\lambda}} \rangle|^p \right)^{1/p} &\lesssim 2^{j(n)(1/r-1/p)} \left( \sum_{\tilde{\lambda} \in \Lambda_n} |f|_{B_{p,p}^{\alpha(1,a_1/a_2)}(\tilde{\Omega}_{\tilde{\lambda}})}^p \right)^{1/p} \\
 &\lesssim 2^{-n\alpha} |f|_{B_{p,p}^{\alpha(1,a_1/a_2)}(\Omega)},
 \end{aligned}$$

i.e.,

$$\|\Omega_n f\|_{p, \Omega} \lesssim 2^{-n\alpha} |f|_{B_{p,p}^{\alpha(1,a_1/a_2)}(\Omega)}, \quad n \in \mathbb{N}.$$

Moreover, we note that

$$\|Q_0 f\|_p = \|P_0 f\|_p \leq \|P_0 f\|_r \lesssim \|f\|_r \lesssim \|f\|_{B_{p,p}^{\alpha(1,a_1/a_2)}}.$$

In the second step we prove that  $S_{p,p}^{\alpha} \hookrightarrow B_{p,p}^{\alpha(1,a_1/a_2)}$ :  $f \in S_{p,p}^{\alpha}$  gives  $f \in L_r(\Omega)$  and, further, that  $f = \sum_{n=0}^{\infty} Q_n f$  in  $L_p(\Omega)$ . Then by Theorem 3.5

$$\|f\|_{B_{p,p}^{\alpha(1,a_1/a_2)}}^p \lesssim \sum_{n=0}^{\infty} \|Q_n f\|_{B_{p,p}^{\alpha(1,a_1/a_2)}}^p \lesssim \sum_{n=0}^{\infty} 2^{p\alpha n} \|Q_n f\|_p^p$$

follows. Finally, interpolation arguments give the result: Choosing appropriate  $\alpha_1 < \alpha < \alpha_2$ , i.e.,  $\alpha = (1 - \theta)\alpha_1 + \theta\alpha_2$  with an  $\theta \in (0, 1)$ , we get by the above arguments that

$$S_{p,p}^{\alpha_i} \hookrightarrow B_{p,p}^{\alpha_i(1,a_1/a_2)} \hookrightarrow S_{p,\infty}^{\alpha_i}.$$

Then interpolation implies for  $0 < q \leq \infty$  that  $S_{p,q}^{\alpha} = (S_{p,p}^{\alpha_1}, S_{p,p}^{\alpha_2})_{\theta,q}$  and  $S_{p,q}^{\alpha} = (S_{p,\infty}^{\alpha_1}, S_{p,\infty}^{\alpha_2})_{\theta,q}$ , and reiteration gives by (38) that  $B_{p,q}^{\alpha(1,a_1/a_2)} = (B_{p,p}^{\alpha_1(1,a_1/a_2)}, B_{p,p}^{\alpha_2(1,a_1/a_2)})_{\theta,q_1}$  thus

$$S_{p,q}^{\alpha} \hookrightarrow B_{p,q}^{\alpha(1,a_1/a_2)} \hookrightarrow S_{p,q}^{\alpha}. \blacksquare$$

### ACKNOWLEDGMENT

This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant Ho 1846/1-1. It was revised and completed while the author held a temporary full position for applied mathematics at the Universität Gesamthochschule Kassel.

### REFERENCES

1. M. Z. Berkolaiko and I. Ya. Novikov, Unconditional bases in spaces of functions of anisotropic smoothness, *Proc. Steklov Inst. Math.* **204**, No. 3 (1994), 27–41.
2. O. V. Besov, Interpolation and embeddings of generalized function spaces  $B_{p,q}^s, F_{p,q}^s$  on a domain, *Proc. Steklov Inst. Math.* **219** (1997), 80–102.
3. J. Brudnyi, A multi-dimensional analogue of a theorem of Whitney, *Mat. Sb.* **11** (1970), 157–170.
4. C. Canuto and A. Tabacco, Multilevel decompositions of functional spaces, *J. Fourier Anal. Appl.* **3** (1997), 715–742.
5. A. Cohen, Wavelets in numerical analysis, in “Handbook of Numerical Analysis,” Vol. VII (P. G. Ciarlet and J. L. Lions, Eds.), Elsevier, North Holland, 1999.
6. W. Dahmen, Multiscale analysis, approximation, and interpolation spaces, in “Approximation Theory. VIII. Wavelets and Multilevel Approximation” (C. K. Chui and L. L. Schumaker, Eds.), pp. 47–88, World Scientific, Singapore, 1995.
7. W. Dahmen, Wavelet and multiscale methods for operator equations, *Acta Numer.* **6** (1997), 55–228.
8. W. Dahmen, R. DeVore, and K. Scherer, Multi-dimensional spline approximation, *SIAM J. Numer. Anal.* **17** (1980), 380–402.
9. W. Dahmen, A. Kunoth, and K. Urban, Biorthogonal wavelets on the interval—Stability and moment conditions, *Appl. Comput. Harmon. Anal.* **6**, No. 2 (1999), 132–196.
10. R. A. DeVore, B. Jawerth, and V. Popov, Compression of wavelet decompositions, *Amer. J. Math.* **114** (1992), 737–785.
11. R. A. DeVore and G. C. Lorentz, “Constructive Approximation,” Springer, Berlin/Heidelberg, 1993.
12. R. A. DeVore and V. Popov, Interpolation of Besov spaces, *Trans. Amer. Math. Soc.* **305** (1988), 397–414.
13. R. A. DeVore and V. Popov, Interpolation spaces and non-linear approximation, in “Function Spaces and Applications” (M. Cwikel *et al.*, Eds.), Lecture Notes in Mathematics, Vol. 1302, pp. 191–205, Springer, Berlin, 1988.
14. R. A. DeVore and R. C. Sharpley, “Maximal Functions Measuring Smoothness,” *Memoirs of the American Mathematical Society*, Vol. 293, AMS, Providence, RI, 1984.
15. W. Farkas, Atomic and subatomic decompositions in anisotropic function spaces, *Math. Nachr.* **209** (2000), 83–113.
16. W. Farkas, J. Johnsen, and W. Sickel, Traces of anisotropic Besov–Triebel–Lizorkin spaces—A complete treatment of the borderline case, *Math. Bohem.* **125**, No. 1 (2000), 1–37.

17. G. Garrigós and A. Tabacco, Wavelet decompositions of anisotropic Besov spaces, preprint, Politecnico di Torino, 1999.
18. R. Hochmuth, Anisotropic Besov spaces and wavelets, in "Proceedings of the International Wavelet Conference, Tanger," pp. 176–179, Institut National de Recherche en Informatique et en Automatique (Hrsg.), Tanger, 1998.
19. R. Hochmuth, "Wavelets in Numerical Analysis and Restricted Nonlinear Approximation," Habilitation-Thesis, Freie Universität Berlin, 1999.
20. R. Hochmuth, Nonlinear anisotropic boundary value problems—Regularity results and multiscale discretizations, *Nonlinear Anal.* **46** (2001), 1–18.
21. R. Hochmuth,  $N$ -term approximation in anisotropic function spaces, preprint, Freie Universität Berlin, 2000.
22. A. Kamont, Isomorphism of some anisotropic Besov and sequence spaces, *Studia Math.* **110**, No. 2 (1994), 169–189.
23. V. I. Kolyada, Rearrangements of functions and embedding theorems, *Russian Math. Surveys* **44**, No. 5 (1989), 73–117.
24. G. C. Kyriazis, Wavelet coefficients measuring smoothness in  $H_p(\mathbb{R}^d)$ , *Appl. Comput. Harmon. Anal.* **3** (1996), 100–119.
25. Y. Meyer, "Ondelettes et opérateurs, I & II," Hermann, Paris, 1990.
26. S. M. Nikol'skij, "Approximation of Functions of Several Variables and Imbedding Theorems," Die Grundlagen der mathematischen Wissenschaften in Einzeldarstellungen, Vol. 205, Springer-Verlag, Berlin, 1975.
27. P. Oswald, "Multilevel Finite Element Approximation," Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.
28. H.-J. Schmeisser and H. Triebel, Anisotropic spaces. I. Interpolation of abstract spaces and function spaces, *Math. Nachr.* **73** (1976), 108–123.
29. H.-J. Schmeisser and H. Triebel, "Topics in Fourier Analysis and Function Spaces," Akademische Verlagsgesellschaft Geest & Porting K.-G., Leipzig, 1987.
30. B. Stöckert and H. Triebel, Decomposition methods for function spaces of  $B_{pq}^s$  and  $F_{pq}^s$  type, *Math. Nachr.* **89** (1979), 247–267.
31. E. A. Storozhenko and P. Oswald, Jackson's theorem in the spaces  $L_p(\mathbb{R}^k)$ ,  $0 < p < 1$ , *Siber. Math. J.* **19** (1978), 630–639.
32. V. N. Temlyakov, "Approximation of Periodic Functions," Nova Science, New York, 1993.
33. H. Triebel, General function spaces. V, *Math. Nachr.* **87** (1979), 129–152.
34. H. Triebel, "Theory of Function Spaces," Monographs in Mathematics, Vol. 78, Birkhäuser-Verlag, Basel, 1983.